Calhoun: The NPS Institutional Archive
DSpace Repository

Theses and Dissertations 1. Thesis and Dissertation Collection, all items

1986-12

On the relative observability of a linear system.

Marques, Antonio Jose Gameiro

http://hdl.handle.net/10945/22012

Copyright is reserved by the copyright owner

Downloaded from NPS Archive: Calhoun

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CALIFORNIA 93943-5002

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

ON THE RELATIVE OBSERVABILITY OF A LINEAR SYSTEM

by

Antonio Jose Gameiro Marques

December 1986

Thesis Advisor D. E. Kirk

Approved for public release; distribution is unlimited.

T232053

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b OFFICE SYMBOL (If applicable) 32	7a NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000	7b ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000
8a NAME OF FUNDING/SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO PROJECT NO TASK NO WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

ON THE RELATIVE OBSERVABILITY OF A LINEAR SYSTEM

12 PERSONAL AUTHOR(S)

Marques, Antonio, G.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Master's Thesis	FROM ______ TO ______	1986 December	91

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Observability, Relative Observability, Linear Systems

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The problem of identification of essential states is considered, leading to the definition of the concept of relative observability. Two analytical methods and two graphical methods to evaluate the relative degree of observability of a linear time-invariant system are presented and three algorithms are derived to rapidly evaluate the relative degree of observability for high order systems. Numerical examples are tested using a computer program. An approach for improving the degree of observability of a specific state variable is proposed for the single output regulator problem, multi-output servo and regulator problems. The improvement algorithm is tested on a single output regulator problem.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION unclassified	
22a NAME OF RESPONSIBLE INDIVIDUAL Donald E. Kirk	22b TELEPHONE (Include Area Code) 408-6242309	22c OFFICE SYMBOL 62Ki

83 APR edition may be used until exhausted
All other editions are obsolete

Approved for public release; distribution is unlimited.

On the Relative Observability of a Linear System

by

Antonio Jose Gameiro Marques
Lieutenant, Portuguese Navy
B.S., Portuguese Naval Academy, 1981

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
December 1986

ABSTRACT

The problem of identification of essential states is considered, leading to the definition of the concept of relative observability. Two analytical methods and two graphical methods to evaluate the relative degree of observability of a linear time-invariant system are presented and three algorithms are derived to rapidly evaluate the relative degree of observability for high order systems. Numerical examples are tested using a computer program. An approach for improving the degree of observability of a specific state variable is proposed for the single output regulator problem, multi-output servo and regulator problems. The improvement algorithm is tested on a single output regulator problem.

M351

TABLE OF CONTENTS

I. INTRODUCTION 8

II. THE CONCEPT OF OBSERVABILITY 10

- A. DEFINITION 10
- B. THE STATE-OBSERVABILITY PROBLEM 10
- C. EXAMPLE 11
  - 1. Observable system 11
  - 2. Non-observable system 11

III. THE CONCEPT OF RELATIVE OBSERVABILITY. 13

- A. DEFINITION 13
- B. METHODS FOR EVALUATING RELATIVE OBSERVABILITY 13
- C. UPPER-BOUND ERROR 14
  - 1. Theoretical background 14
  - 2. Example 17
- D. STANDARD DEVIATION ERROR CRITERION 18
  - 1. Theoretical background 18
  - 2. Example 19
- E. GRAPHICAL METHOD 20
- F. A MORE GENERALIZED METHOD 29
- G. SUMMARY 33

IV. IMPROVEMENT OF THE DEGREE OF OBSERVABILITY OF A WEAKLY OBSERVABLE SYSTEM 34

- A. INTRODUCTION 34
- B. THEORETICAL BACKGROUND 34
- C. DERIVATION OF THE METHOD 35
- D. THE MULTI-OUTPUT CASE 45
- E. CASE WHERE THERE IS A CONTROL INPUT 55
- F. SUMMARY 58

V. ANALYSIS OF THE RESULTS 60
A. INTRODUCTION 60
B. PERFORMANCE OF TOBS 60
C. PERFORMANCE OF COBS 66
D. SUMMARY 72

VI. CONCLUSION 73

APPENDIX A: TESTING THE DEGREE OF OBSERVABILITY OF AN LTI SYSTEM 76

APPENDIX B: IMPROVING THE DEG. OF OBS. OF A WEAKLY OBSERVABLE SYSTEM 80

LIST OF REFERENCES 88

INITIAL DISTRIBUTION LIST 90

LIST OF TABLES

1. TOBS OUTPUT FOR ABLIN'S SECOND ORDER EXAMPLE 60
2. SUMMARY OF THE RESULTS OBTAINED USING THREE DIFFERENT METHODS 62
3. TOBS OUTPUT FOR A SIXTH ORDER EXAMPLE 64
4. TOBS OUTPUT FOR THE SL-7 HULL MODEL -CASE I- 65
5. TOBS OUTPUT FOR THE SL-7 HULL MODEL -CASE II- 66
6. TOBS OUTPUT FOR THE MODEL OF EXAMPLE 2 OF [REF. 19] 67
7. NEW RANKING OF THE STATE VARIABLES FOR ABLIN'S EXAMPLE 68

LIST OF FIGURES

3.1 Graphical interpretation of ill-conditioned systems-I 22
3.2 Graphical interpretation of ill-conditioned systems-II 24
3.3 Case of different degrees of observability 26
3.4 Relation between the row entries of Q^T and the vectors v_1 and v_2 28
3.5 Importance of the angles on evaluating the degree of observability per state-variable 30
4.1 Algorithm to provide the minimum of the condition number of Q^T by changing a specific entry in C 46
5.1 Graphical interpretation of Ablin's example before improvement 70
5.2 Graphical interpretation of Ablin's example after improvement 71

I. INTRODUCTION

The concept of observability introduced by Kalman in 1960 [Ref. 1], is, along with the dual concept of controllability, one of the most significant ideas to come out of the state space approach to control theory. The concept of observability presented in most of the literature available today, gives only an answer to whether the system is observable or not. However, another question was raised by Brown in 1966 [Ref. 2], opening the issue of how observable a given system is. Are all the states equally observable? If not, what are the relative degrees of observability of the state variables? Since 1966, this issue has been investigated in [Ref. 3], [Ref. 4], [Ref. 5] and [Ref. 6] among others. In the majority of theses investigations the goal has been either the state-space reduction of the model or frequency domain model reduction. Silverman has probably been one of the most active researchers on this subject, using, in several of his papers, the technique of rotating the coordinate system in such a way that a so called "balanced realization" is obtained. By studying this new realization the "nearly redundant" states can be distinguished from the most important ones, providing the required information to reduce the order of the model.

In this thesis the problem of how observable the states are is again studied using techniques different from the ones mentioned above. Two graphical methods are developed: the first one is relatively simple, but rather helpful in understanding the concepts involved in ranking the state variables according to their relative degrees of observability. The second graphical method, based on some ideas given in [Ref. 7], is more elaborate. Both graphical methods are compared against a third method described in [Ref. 7] whose development is explained in this thesis.

The tests to find the relative observability for each state variable provide the information to develop a method that enables the improvement of the degree of observability of the least observable state. Although all the basic derivations assume a single output regulator problem, the generalization to multi-output non-regulator systems is also made. This thesis is, therefore, divided as follows: in Chapter II a brief and general definition of observability is given followed by two single output numerical examples. Chapter III begins with a definition of relative observability followed by the presentation of methods to evaluate the relative degree of observability of a system.

Three algorithms are presented leading to a Fortran computer program whose listing is in Appendix A. In Chapter IV an algorithm to improve the relative degree of observability of a weakly observable state is developed for the single output regulator problem. This algorithm leads to a Fortran program that is contained in Appendix B. The generalization for multi-output servo systems is also discussed. Finally in Chapter V the performances of the three methods of Chapter III are compared as well as the performance of the algorithm developed in Chapter IV. Some particular examples used with the computer programs are presented and analized in some detail.

II. THE CONCEPT OF OBSERVABILITY

A. DEFINITION

A system is said to be completely observable if, from the measurements of the output y(t) over a finite period of time, $t \in [0, t_0]$, every initial state can be determined [Ref. 8: p. 764], [Ref. 9: p. 107]. In other words, if every transition of the states affects every element of the output vector, the system is completely observable [Ref. 8: p. 764].

B. THE STATE-OBSERVABILITY PROBLEM

Consider a linear time-invariant system described by equations 2.1 and 2.2

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2.1}$$

$$y(t) = Cx(t) \tag{2.2}$$

where:

x(t) is a nx1 column vector of the state-variables.

$\dot{x}(t)$ is a nx1 column vector of the time derivatives of the state variables

A is a nxn matrix.

B is a nxp matrix

u is a px1 column vector of the inputs to the system.

y(t) is a mx1 column vector of the output variables of the system

C is a mxn matrix

The effect of the input signal u(t) can always be determined, if x(0) can be computed from the measurements of the output, i.e., if the system is completely observable. Therefore, there is no loss of generality by assuming that u(t) = 0 [Ref. 9: p. 107]. Under this assumption, to evaluate a system's observability a matrix Q, called the observability matrix, must be formed. [Ref. 10: p. 108].

$$Q = [C^T \quad A^T C^T \quad (A^T)^2 C^T \quad \ldots \quad (A^T)^{n-1} C^T] \tag{2.3}$$

Each of the submatrices separated by spaces in equation 2.3 is nx1. Therefore, there will be n columns in the Q matrix; this matrix will be square (nxn) when there is only a single output. If the n columns of Q are all linearly independent, i.e., if the nxn matrix Q is nonsingular the initial state vector x(0) can always be found from the measurements of the output [Ref. 10: p. 105]. On the other hand, if Q is singular, (not of full rank), some columns of Q will be linearly dependent on others, and only special choices of y(t) will produce the required solution [Ref. 10: p. 105].

C. EXAMPLE

1. Observable system

Consider the single output linear time invariant system described by:

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y = Cx(t)$$

Where:

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$

The matrix Q is equal to:

$$Q = \begin{bmatrix} 1 & 1 & 7 \\ 0 & 3 & 12 \\ 1 & 1 & 1 \end{bmatrix}$$

This matrix Q is nonsingular, i.e., of full rank, therefore the system is observable.

2. Non-observable system

Consider, again the same state representation where the matrices A, B and C are as follows:

$$A = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

The matrix Q is equal to:

$$Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 1 & 1 & -3 \end{bmatrix}$$

In this case the matrix Q is singular, i.e., its determinant is equal to zero. Its rank is also less than three, (actually two), because the maximum submatrix of Q with determinant different from zero is a 2x2 matrix. Therefore, in this case there is one state that is not observable. The identification of that non observable state is not possible just by performing the previously described test, where the answer "yes-no" is given, with no indication of how observable the system is and which are the most observable states. In the following Chapters methods will be discussed and presented where answers to these questions will be given.

III. THE CONCEPT OF RELATIVE OBSERVABILITY.

A. DEFINITION

The problem of relative observability deals with finding an algorithm or figure of merit for each state-variable, which will reflect how observable the state-variable is. Measuring $y(t)$ at times $t = t_1, t_2, \ldots, t_n$, one wants to be able to evaluate which state components can be most accurately determined from the n set of measurements of the output. In other words, how an error in the measurements is reflected in determining the initial states? If a small error in the measurements produces a large error in the calculated state, this state is, therefore, difficult to determine from the measurement of the output. The output has a "fuzzy image" of the state that is "hidden" by the large error. If a small error in the measurements produces a small error in the evaluation of a particular state, this state will be evaluated more accurately than the others from the measurements of the output. Due to the small error the "image" of the state at the output is less "fuzzy".

B. METHODS FOR EVALUATING RELATIVE OBSERVABILITY

The three methods that will be described for evaluating the relative degree of observability were developed based on the concept given previously: the projection of a measurement error in computing the initial state.

Starting with equations 2.1 and 2.2 and assuming the input to be zero, equation 2.2 is differentiated n-1 times and equation 2.1 is substituted as shown below [Ref. 10: p. 106].

$$
\begin{aligned}
y(t_0) &= Cx(t_0) \\
\dot{y}(t_0) &= C\dot{x}(t_0) = CAx(t_0) \\
\ddot{y}(t_0) &= CA\dot{x}(t_0) = CA^2x(t_0) \\
&\;\;\vdots \qquad\quad \vdots \qquad\quad \vdots \\
\overset{(n-1)}{y}(t_0) &= CA^{n-1}x(t_0)
\end{aligned}
\tag{3.1}
$$

The result is n equations with n unknowns. It can be seen that the coefficient vectors C, CA, $CA^2, \ldots, CA^{(n-1)}$ are row vectors with the same elements as the respective column vectors C^T, $A^TC^T, (A^T)^2C^T, \ldots, (A^T)^{n-1}C^T$ of the Q matrix. Thus, the elements of the columns of the Q matrix are the coefficients of the equations relating the initial states with the measured quantities. Equation 3.1 can be written as

$$y_d(t_0) = Q^T x(t_0) \tag{3.2}$$

where $y_d(t_0)$ is the column vector of the output and its n-1 derivatives evaluated at $t = t_0$.

The fact that the output equation 2.2 must be differentiated n-1 times to obtain n equations and n unknowns, implies that the output has to be measured n times. If the system has only one output, as in this case, the state vector, $x(t_0)$, can be found by inverting the Q^T matrix. The inverse of Q^T only exists if its determinant is nonzero or, in other words, if the rank of the matrix is equal to its order. However, if the coefficients of any pair of equations are approximately proportional, the determinant of Q^T will be a small number and, from a practical point of view, it will be difficult to invert the matrix and solve the system of equations. A matrix that has a small number for its determinant is an "ill-conditioned" matrix. Since a zero value for the determinant occurs only for a singular matrix, it can be said that an ill-conditioned matrix is nearly singular.

The solutions of ill conditioned systems are very sensitive to small changes in either the coefficients of the matrix or the values of the forcing functions. Therefore, if the rows of Q^T are nearly proportional, a small error in in the measurement of $y_d(t_0)$, will produce a large error in $x(t_0)$ due to the high sensitivity of ill-conditioned systems.

C. UPPER-BOUND ERROR

1. Theoretical background

The degree of independency of the rows of Q^T can be used to quantify the degree of observability for each state variable [Ref. 7: p. 12]. If two or more of the rows of Q^T are almost linearly dependent the entries of $(Q^T)^{-1}$ will be large numbers. In defining the degree of observability for each state variable using the upper-bound error, the reciprocal of the increase in the error of the calculated state-variable over the observation error will be used [Ref. 7: p. 17]. The reciprocal will be used because a small number associated with a low degree of observability is more intuitively appealing than a large number.

Referring to matrix equation 3.2, normalization of the rows of Q^T will be performed. The goal of this operation is to obtain a maximum value for the degree of observability equal to one. To accomplish the normalization the elements of $y_d(t_0)$ are divided by the length of the corresponding row vector of the Q^T matrix. The normalized vector is defined as $y_n(t_0)$ and the normalized Q^T as P. Equation 3.2 can now be written as :

$$y_n(t_0) = Px(t_0) \tag{3.3}$$

The vector $y_n(t_0)$ consists of the actual value of the measurements $y_a(t_0)$ and an error e_m, thus, equation 3.3 can be written as :

$$y_a(t_0) + e_m = Px(t_0) \tag{3.4}$$

Solving equation 3.4 for $x(t_0)$ by taking the inverse of P results in

$$x(t_0) = P^{-1}y_a(t_0) + P^{-1}e_m \tag{3.5}$$

Where:

$x(t_0)$ is the nx1 vector of the computed states at time t_0.
$P^{-1}y_a(t_0)$ is the nx1 vector of the true states.
$P^{-1}e_m$ is the nx1 vector of the error in the computed states.

Defining

$$x_a(t_0) = P^{-1}y_a(t_0) \tag{3.6}$$

and substituting this in equation 3.5 yields

$$x(t_0) = x_a(t_0) + P^{-1}e_m \tag{3.7}$$

or

$$x(t_0)-x_a(t_0)=P^{-1}e_m \tag{3.8}$$

The left hand side of equation 3.8 is actually the error in the computed state, e_c. Substituting this notation in equation 3.8 gives :

$$e_c=P^{-1}e_m \tag{3.9}$$

Recalling what has been said about the dependency among the rows of P, it was stated that if the dependency was high the system was ill-conditioned and the outcome was large numbers for the rows of P^{-1}. Equation 3.9 shows that the error in the computed states is a linear combination of the measurement errors. For equal values of error in the measurements, i.e., $e_{1m}=e_{2m}=\ldots=e_{nm}$, the entries in the e_c vector will be different according to the magnitudes of the entries in the corresponding rows of P. The larger the sum of magnitudes of the entries in a row, the larger will an error in $y_d(t_0)$ be reflected in $x(t_0)$.

To recapitulate , the upper-bound error criterion was defined as the reciprocal of the ratio of the error in the calculated state-variable to the observation error.

$$\delta_{obsi}=\frac{1}{\frac{e_{ic}}{e_{im}}}=\frac{e_{im}}{e_{ic}} \tag{3.10}$$

Where:

δ_{obsi} is the degree of observability of the ith state variable.

e_{im} is the error in the ith measurement.

e_{ic} is the error in computing the ith state.

Equation 3.9 can be rewritten as equation 3.11, where the α_{ij} are the entries of P^{-1}.

$$\begin{aligned}
e_{1c}&=\alpha_{11}e_{1m}+\alpha_{12}e_{2m}+\ldots+\alpha_{1n}e_{nm}\\
e_{2c}&=\alpha_{21}e_{1m}+\alpha_{22}e_{2m}+\ldots+\alpha_{2n}e_{nm}\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\\
e_{nc}&=\alpha_{n1}e_{1m}+\alpha_{n2}e_{2m}+\ldots+\alpha_{nn}e_{nm}
\end{aligned} \tag{3.11}$$

Substituting equation 3.11 in 3.10 yields

$$\delta_{obsi} = \frac{e_{im}}{(\alpha_{i1}e_{1m} + \alpha_{i2}e_{2m} + \ldots\ldots + \alpha_{in}e_{nm})} \tag{3.12}$$

The main goal is to obtain a measure of the magnitude of the error in the computed states, for an equal error made while measuring the output. Therefore, it is reasonable to consider as equal all of the e_m. If one does not do so, misleading values for δ_{obs} for the state-variables may be obtained, because the comparison between δ_{obs1}, δ_{obs2}, ..., δ_{obsn}, will be done under different conditions.

Consider, then, that all the errors in the measurements are equal. Factoring out the e_m in the denominator of equation 3.12 and cancelling this value with the numerator, yields equation 3.13 :

$$\delta_{obsi} = \frac{1}{\sum (\alpha_{ij})} \tag{3.13}$$

Since one is looking for the maximum possible value for the error in computing the ith state, the signs of the elements in the e_m vector must be such that an upper-bound for the computed error e_{ic} will be obtained. Since the computed error is a linear combination of the observation errors, the upper-bound for that error will be given by the sum of the absolute values of the row entries of P^{-1}. This is expressed by

$$\delta_{obsi} = \frac{1}{\sum |\alpha_{ij}|} \tag{3.14}$$

which is the final equation for evaluating the degree of observability for each state-variable using the upper-bound error method.

2. **Example**

Consider the single output linear time-invariant system described by:

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$

Where:

$$A = \begin{bmatrix} 0 & 1 \\ -4 & -5 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

STEP ONE - Form the Q matrix:

$$Q = \quad [C^T \; A^T C^T] \quad => \quad Q = \begin{bmatrix} 0 & -4 \\ 1 & -5 \end{bmatrix}$$

STEP TWO - Form Q^T and normalize its rows, producing the matrix P:

$$Q^T = \begin{bmatrix} 0 & 1 \\ -4 & -5 \end{bmatrix} \quad => \quad P = \begin{bmatrix} 0 & 1 \\ -0.625 & -0.781 \end{bmatrix}$$

STEP THREE - Find P^{-1}:

$$P^{-1} = \begin{bmatrix} -1.249 & -1.6 \\ 1 & 0 \end{bmatrix}$$

STEP FOUR - Find the degree of observability using equation 3.13:

$$\delta_{obs1} = \frac{1}{|-1.249| + |-1.6|} = 0.351$$

$$\delta_{obs2} = \frac{1}{1 \; + \; 0} = 1$$

The upper-bound error criterion for x_1 says that the error on computing $x_1(t_0)$ will be 2.85 times bigger than the errors made while measuring $y(t_0)$ and its derivative if the assumption that both measurement errors are equal holds. The upper-bound error method criterion for x_2 tells the user that the error on computing $x_2(t_0)$ will be the same as the one made on the measurements. This result is very satisfying because, looking at the C matrix, one can see that the state x_2 is measured directly while x_1 influences the output measurement indirectly.

D. STANDARD DEVIATION ERROR CRITERION

1. Theoretical background

Assume now that the errors in the output measurements that are being made are statistically independent random variables whose second order statistics are known. It is known that the variance of a linear combination of statistically independent random variables is the sum of the coefficients squared, multiplied by the respective variances of each random variable [Ref. 11: p. 100]. Thus, taking the variance of the ith measurement and dividing it by the variance of the error in computing the ith state-variable gives

$$\delta_{obsi} = \frac{\mathrm{Var}[e_{im}]}{\mathrm{Var}[\alpha_{i1}e_{1m} + \alpha_{i2}e_{2m} + \ldots\ldots + \alpha_{in}e_{nm}]} \qquad (3.15)$$

If the stated assumption holds, i.e., if the measurements of the output are statistically independent random variables, the following equation is obtained:

$$\delta_{obsi} = \frac{\mathrm{Var}[e_{im}]}{\alpha_{i1}^2\ \mathrm{Var}[e_{1m}] + \alpha_{i2}^2\ \mathrm{Var}[e_{2m}] + \ldots\ldots + \alpha_{in}^2\ \mathrm{Var}[e_{nm}]} \qquad (3.16)$$

Here again the assumption of equal variance is made because this is the only way of comparing the n degrees of observability under the same conditions. Factoring out the variances in equation 3.16 and cancelling with the numerator results in equation 3.17

$$\delta_{obsi} = \frac{1}{\alpha_{i1}^2 + \alpha_{i2}^2 + \ldots\ldots + \alpha_{in}^2} \qquad (3.17)$$

$$\delta_{obsi} = \frac{1}{\sum |\alpha_{ij}|^2} \qquad (3.18)$$

Equation 3.18 is the final equation for evaluating the degree of observability for each state-variable using the variance as a measure of the error.

2. **Example**

Starting in step three of the previous example:

$$P^{-1} = \begin{bmatrix} -1.249 & -1.6 \\ 1 & 0 \end{bmatrix}$$

Find the degree of observability using equation 3.18:

$$\delta_{obs1} = \frac{1}{(-1.249)^2 + (-1.6)^2} = 0.243$$

$$\delta_{obs2} = \frac{1}{(1)^2 + 0} = 1$$

The variance criterion for evaluating the degree of observability shows that the error in computing $x_1(t_0)$ has a variance 4.12 times bigger than the one encountered in the measurements. The variance of the error in evaluating $x_2(t_0)$ is equal to the variance of the error in the measurements.

E. GRAPHICAL METHOD

Another interpretation of ill conditioning is obtained by examining graphically a system of two equations and two unknowns. Consider the following system:

$$Q^T x(t_0) = y(t_0) \tag{3.19}$$

where:

$$Q^T = \begin{bmatrix} 1.01 & 0.99 \\ 0.99 & 1.01 \end{bmatrix}$$

Substituting the matrix Q^T in equation 3.19 yields

$$1.01x_1(t_0) + 0.99x_2(t_0) = y_1(t_0) \tag{3.20}$$

$$0.99x_1(t_0) + 1.01x_2(t_0) = y_2(t_0) \tag{3.21}$$

Equations 3.20 and 3.21 can be seen as the equations of two lines with the abcissa $x_1(t_0)$, the ordinate $x_2(t_0)$ and the $x_2(t_0)$ intercepts at $x_1(t_0)=0$ being proportional to $y_1(t_0)$ and $y_2(t_0)$ respectively. The measurements of the output are $y_1(t_0)$ and $y_2(t_0)$ and the main interest is to evaluate how much an error or disturbance made in the measurements (e_m) is "amplified" when the computation of the initial states takes place. In Figure 3.1 the lines corresponding to equations 3.20 and 3.21 are presented. The lines intersect in the point $x_{1a}(t_0)$, $x_{2a}(t_0)$ which is the actual solution of the system of equations. Suppose that some errors are made while measuring the outputs $y_1(t_0)$ and $y_2(t_0)$. How is this fact going to change the position of the original lines? How is the solution of the original system going to be perturbed? As in the development and discussion of Ablin's method [Ref. 7], the errors in the measurements will be considered as equal so that the comparison between the different degrees of observability will be done under the same conditions. Consider, then, that the value measured for $y_1(t_0)$ exceeds its real value by an amount that is equal to e_m. In this case the line corresponding to equation 3.20 is displaced in the direction of positive $x_2(t_0)$ and $x_1(t_0)$ to a position parallel to its original one. Conversely, if the measured value is less than its real one by an amount that is equal to e_m, the line corresponding to equation 3.20 is displaced to a position parallel to its original one but now in the direction of the negative $x_2(t_0)$ and $x_1(t_0)$. These two lines form an uncertainty region around the line corresponding to equation 3.20. The same thing occurs for the line corresponding to equation 3.21 and a second uncertainty region is obtained. As can be seen in Figure 3.1 the intersection of these two regions forms a parallelogram shaped region that contains all the possible solutions of the system of equations if the measurements can vary between $y_1(t_0)+e_m$ and $y_1(t_0)-e_m$ for the first measurement of the output and $y_2(t_0)+e_m$ and $y_2(t_0)-e_m$ for the second one. The boundaries of the parallelogram shaped region contain the solution of the system of equations for the particular cases where the errors in the measurements are maximum in magnitude. In Figure 3.1 eight different points are marked on the boundary region. If the errors in the measurements are such that $y_1(t_0)$ is measured at its minimum allowable bound, namely $y_1(t_0)-e_m$, and $y_2(t_0)$ is measured at its maximum allowable bound, point 1 is obtained. In the case of point 2 no error was assumed in the measurement of $y_1(t_0)$ while the other measurement was corrupted by $+e_m$. Point 3 depicts the case where both measurements were affected by the same positive error. In the case of point 4, $y_2(t_0)$ was measured with no error but the error in $y_1(t_0)$ is $+e_m$. Point 5 is, in a way, symmetric to point 1: $y_1(t_0)$ is measured corrupted by its maximum allowable error while $y_2(t_0)$ is measured with its minimum allowable bound. In the case of point 6, no error is assumed while measuring $y_1(t_0)$, but the second measurement of the output is affected by an error equal to $-e_m$. Point 7 is symmetric to point 3: both measurements of the output are affected by the same negative error $-e_m$. Finally, point 8 is the case where no error was present in the measurement of $y_2(t_0)$ but $y_1(t_0)$ was affected by an error equal to $-e_m$. In all cases the solution of the system of equations is given by the projection of these points onto the axes $x_1(t_0)$ and $x_2(t_0)$.

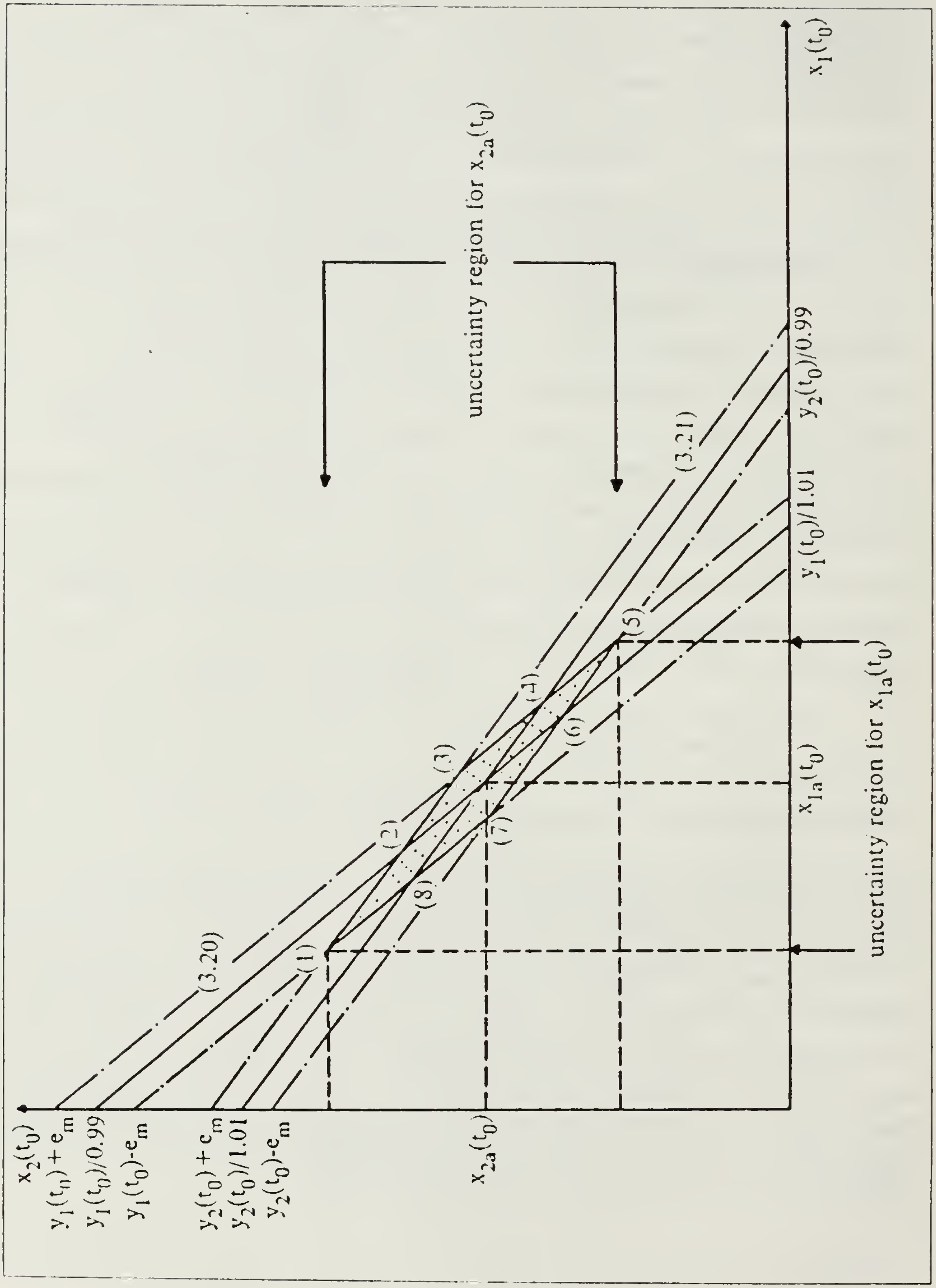

Figure 3.1 Graphical interpretation of ill-conditioned systems-I.

The discussion above gives enough information to answer the question of how the true solution of the system of equations is perturbed when errors in the measurements exist. As seen in Figure 3.1, the projections of points 1 and 5 onto the axes yield the most perturbed solutions. On the other hand, the projection of points 3 and 7 yield solutions that are as close as possible to the true solution, assuming that, if any errors are made in the measurements, they are equal to $\pm e_m$. Therefore, if the errors in the measurements are of opposite signs and equal magnitudes, the most perturbed solutions are obtained. In other words, if the errors in the measurements are as far as they can possibly be from one another the corresponding solution of the system of equations is also as far as it can be from the true solution. Conversely, if the errors in the measurements are equal both in magnitude and in sign, the least perturbed solution for this maximum error case is obtained. Thus, points 1 and 5 correspond to the worst possible case when the errors corrupting the measurements are maximum while points 3 and 7 correspond to the best possible case when the maximum errors in the measurements are present. The projection of points number 1 and 5 onto the axes $x_1(t_0)$ and $x_2(t_0)$ gives the maximum possible uncertainty region for the solution of the system of equations. As can be seen in Figure 3.1, these uncertainty regions are of the same order of magnitude for both states but much larger than the errors made in the measurements.

Suppose that the Q^T matrix produces lines that are perpendicular. As can be seen in Figure 3.2 the points 1, 5, 3 and 7 produce solutions that are nearly equally disturbed. Furthermore, the uncertainty regions in the solution of the system of equations are of the same order of magnitude as the errors in the measurements. Therefore, for the cases where the lines describing the system are orthogonal, the uncertainty regions in the solution of the system of equations are of the same order of magnitude as the errors made while measuring the output.

Notice the similarities and differences between the cases depicted in Figures 3.1 and 3.2. In both cases the maximum errors in the measurements are assumed to be equal. However, in the case of Figure 3.1 the errors in the measurements are enormously amplified when the computation of the states takes place, while in the case of Figure 3.2 the amplification is practically nonexistent. In both cases the uncertainty regions for both state variables are nearly equal yielding the conclusion that both state variables are equally observable.

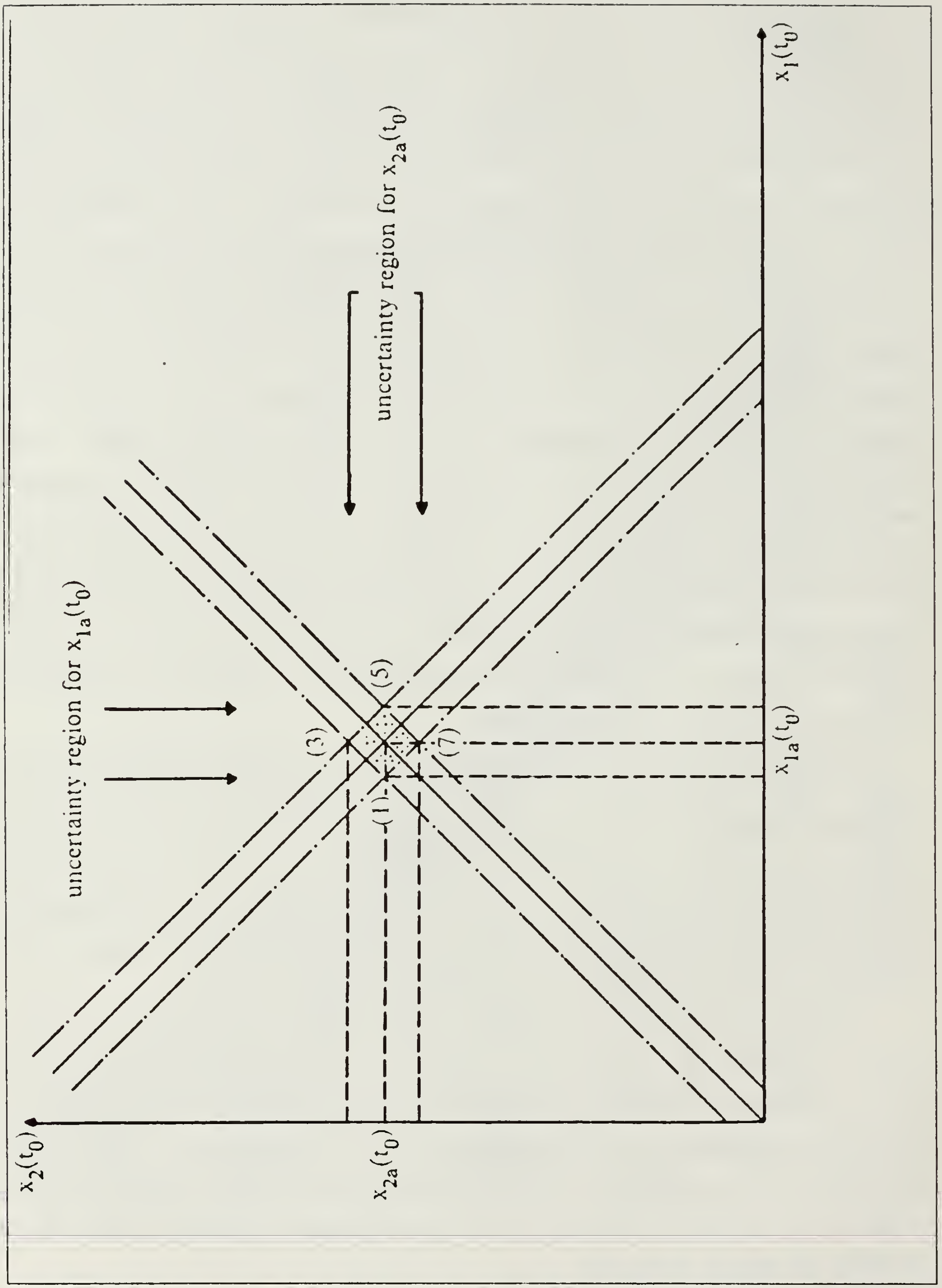
uncertainty region for $x_{2a}(t_0)$
$x_1(t_0)$
uncertainty region for $x_{1a}(t_0)$
(5)
(3)
(7)
(1)
$x_{1a}(t_0)$
$x_{2a}(t_0)$
$x_2(t_0)$

Figure 3.2 Graphical interpretation of ill-conditioned systems-II.

What is behind the difference of behaviour in these two cases? One thing that rapidly comes to mind is the relative position of the lines describing the system under study. In the case depicted in Figure 3.1, these lines have almost the same slope, i.e., they are nearly parallel. In the case of Figure 3.2 the lines are orthogonal.

In the cases studied so far both state variables had approximately the same degree of observability. What happens if the states have different degrees of observability? Figure 3.3 depicts such a case. Suppose that the Q^T matrix produces lines whose angles with the $x_1(t_0)$ axes are very small when compared with the angles between the lines and the $x_2(t_0)$ axis. In this case the projection of the points 1 and 5 onto the axes $x_1(t_0)$ produces an uncertainty region that is much larger than the errors in the measurements. The projection of the same points onto the $x_2(t_0)$ axes produces an uncertainty region that is much smaller than the one obtained for the other state variable and it is also of about the same order of magnitude as the error in the measurement. Thus, the same error made in the measurements is enlarged a substantial amount when the computation of $x_1(t_0)$ takes place while it remains unchanged when $x_2(t_0)$ is computed. Therefore, according to the definition of relative observability one can say that state $x_2(t_0)$ is relatively more observable than state $x_1(t_0)$. Notice, however, that if by any chance the errors in the measurements are maximum but with equal sign (points 3 and 7), the uncertainty region for $x_1(t_0)$ is very small while, for $x_2(t_0)$, the uncertainty region doesn't differ much from the case of points 1 and 5. Therefore, the sensitivity of the solution to the variation of the errors is also meaningful insofar as the relative observability is concerned: if the state is weakly observable it is also more sensitive to changes in the errors of the measurements. If the state is strongly observable it is not very sensitive to changes in the errors of the measurements.

The three cases described so far revealed two kinds of information: one imbedded in the relative position of the lines describing the system and the other imbedded in the position of these lines with respect to the coordinate axis. The first one determines the order of magnitude of the uncertainty regions of the computed state variables: lines nearly orthogonal imply uncertainty regions for the states of the same order of magnitude as the uncertainty regions for the measurements. Lines nearly parallel imply the possibility of at least one uncertainty region for the states much larger than the uncertainty region for the measurements.

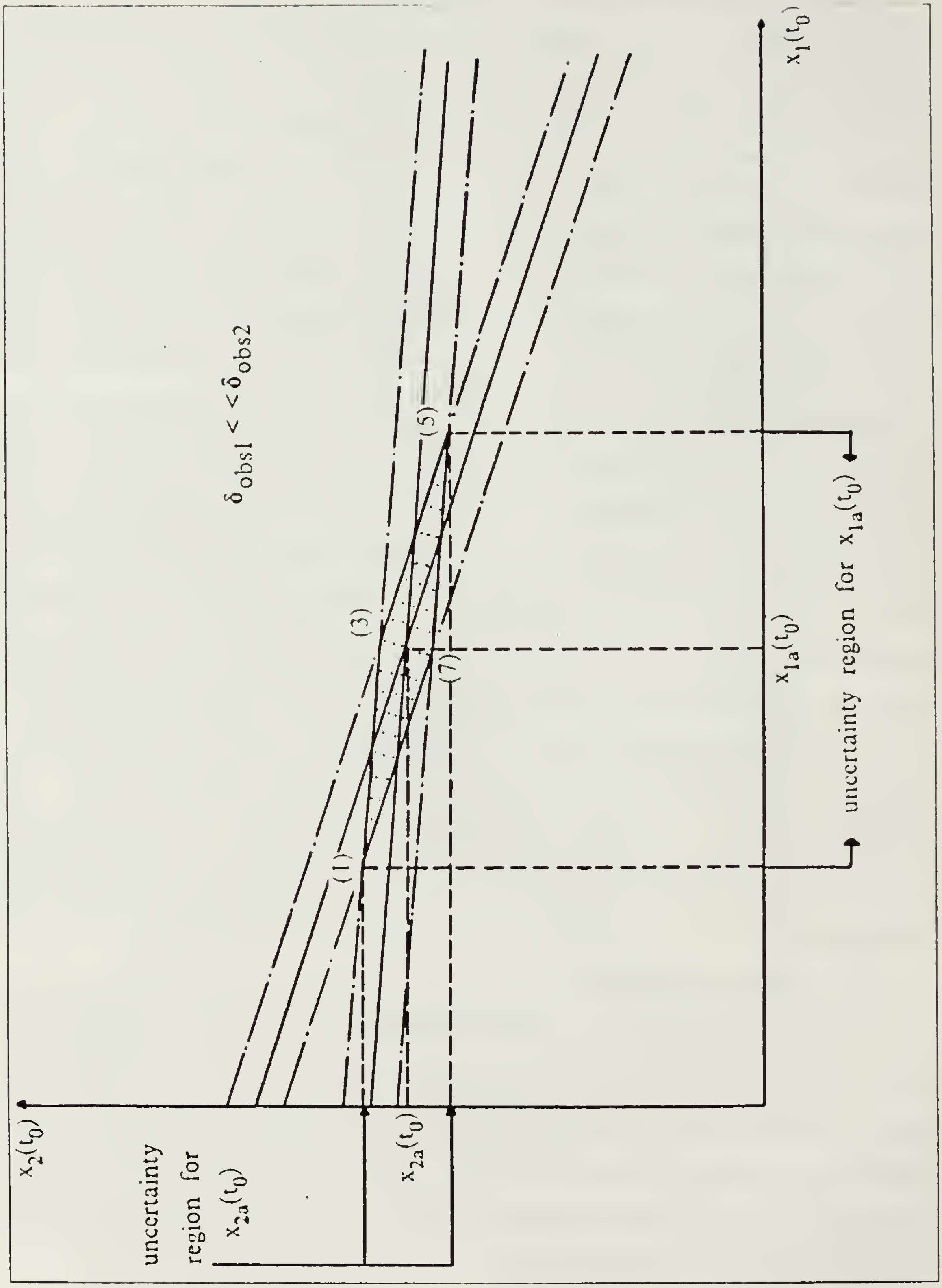
$x_1(t_0)$
$x_2(t_0)$
$\delta_{obs1} << \delta_{obs2}$
(1)
(3)
(5)
(7)
$x_{1a}(t_0)$
uncertainty region for $x_{1a}(t_0)$
$x_{2a}(t_0)$
uncertainty region for $x_{2a}(t_0)$

Figure 3.3 Case of different degrees of observability.

The second information is the more important because it seems to be directly related to the evaluation of the relative degree of observability for each state variable. In the cases depicted in Figures 3.1 and 3.3 the relative positions of the lines were nearly the same. However, their positions with respect to the coordinate axes were not: in the case depicted in Figure 3.1 the angles between the lines and both axes were about the same while in the case depicted in Figure 3.3 the angles between the lines and $x_1(t_0)$ were very small and the angles between $x_2(t_0)$ were almost 90 degrees. As previously discussed, in the case of Figure 3.1 both state variables were equally weakly observable. In the case of Figure 3.3 state variable $x_1(t_0)$ was weakly observable while $x_2(t_0)$ was relatively strongly observable. Thus, the angles between the lines and the coordinate axes are strongly related to the relative degree of observability of the state variables: if these angles are small (lines nearly parallel to the axes), the corresponding state variable is weakly observable. If the angles are nearly 90 degrees the corresponding state variable is strongly observable.

Figure 3.4 shows the case depicted in Figure 3.1 deprived of several auxiliary lines to enable a better understanding of the discussion to follow. Consider the lines perpendicular to those describing the system: the equation of the line orthogonal to the line described by equation 3.20 that passes through the origin is given by :

$$1.01x_2(t_0) = 0.99x_1(t_0) \tag{3.22}$$

The equation of the line orthogonal to the line described by equation 3.21 that passes through the origin is given by :

$$1.01x_1(t_0) = 0.99x_2(t_0) \tag{3.23}$$

Consider, now, the unit vectors u_1 and u_2 such that $u_1 \in x_1(t_0)$ and $u_2 \in x_2(t_0)$. As can be seen in Figure 3.4 , the vectors v_1 and v_2 belong respectively to the lines described by equations 3.22 and 3.23 and are given by:

$$v_1 = 1.01u_1 + 0.99u_2 \tag{3.24}$$

and

$$v_2 = 0.99u_1 + 1.01u_2 \tag{3.25}$$

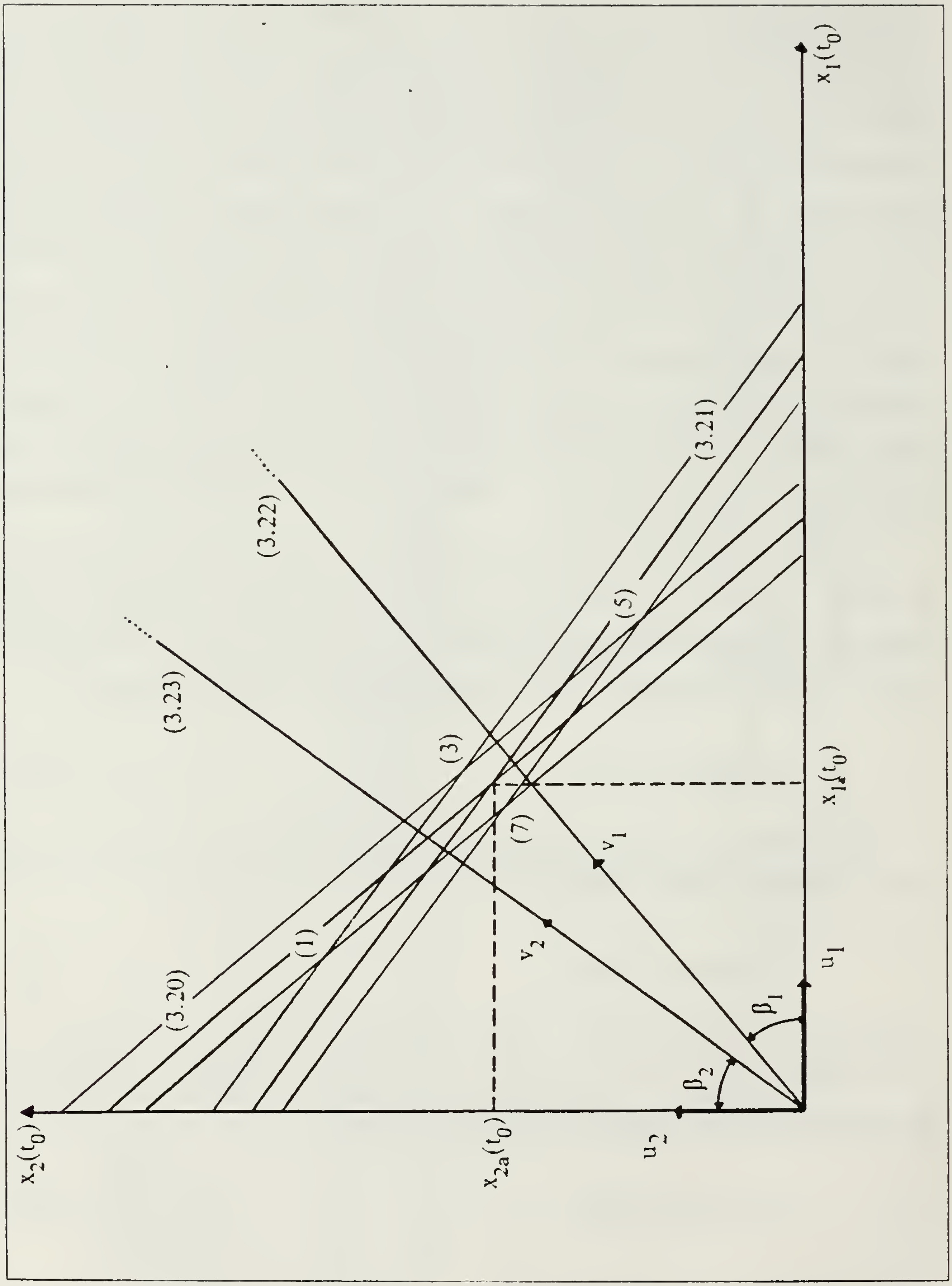

Figure 3.4 Relation between the row entries of Q^T and the vectors v_1 and v_2.

The components of the vectors v_1 and v_2 are the entries in the first and second rows of Q^T or the first and second columns of Q. As can be seen in Figure 3.4 the angles between the vectors v_1 and v_2 and the coordinate axes are nearly equal, i.e., $\beta_1 \sim \beta_2$. This Figure depicts the case where both state variables have equal degrees of observability. Thus, equal angles between the vectors, whose components are the row entries of the Q^T matrix, and the coordinate axes may imply that both state variables are equally observable.

Figure 3.5 shows the case depicted in Figure 3.3 where state variable $x_2(t_0)$ was relatively more observable than state variable $x_1(t_0)$. In this case both v_1 and v_2 are "closer" to $x_2(t_0)$, i.e., $\theta_1 < \theta_2$ and $\gamma_1 < \gamma_2$.

Analyzing and comparing the two cases just discussed, the importance of the information imbedded in the angles between the vectors, (whose components are the row entries of the Q^T matrix), and $x_1(t_0)$ and $x_2(t_0)$, can be understood: the smaller the angle between the normal vectors v_1 and v_2 and an axis of the state-space, the higher the accuracy of computing the state corresponding to that axis, i.e, the higher the degree of observability for that state variable. The angles between the vectors and the coordinate axes are directly related to the magnitudes of their components along these axes: the larger the magnitude of the component along a particular axes, the smaller the angle between the vector and the axis. As previously discussed, the smaller the angle the higher the accuracy when computing that state. Recalling that the rows of Q^T are the vectors relating $x_1(t_0)$ and $x_2(t_0)$ with $y_1(t_0)$ and $y_2(t_0)$, an easy way to check the relative observability of a state variable is to look for the largest principal axis component of each normal vector and then associate these components with the state variables. This idea has been used as an algorithm for the evaluation of the relative degree of observability of a system and is one of the three algorithms presented in the Fortran program TOBS in Appendix A.

F. A MORE GENERALIZED METHOD

The graphical method previously described is quite helpful for low-order systems because it requires few operations, is rapid and gives a good feel for the rank of the state variables as far as relative observability is concerned. When the system becomes larger and more complex the direction of minimum error, although perpendicular to the direction of maximum error, may be difficult to find due to the increasing number of vectors in a higher order state-space. What if instead of checking the whole space

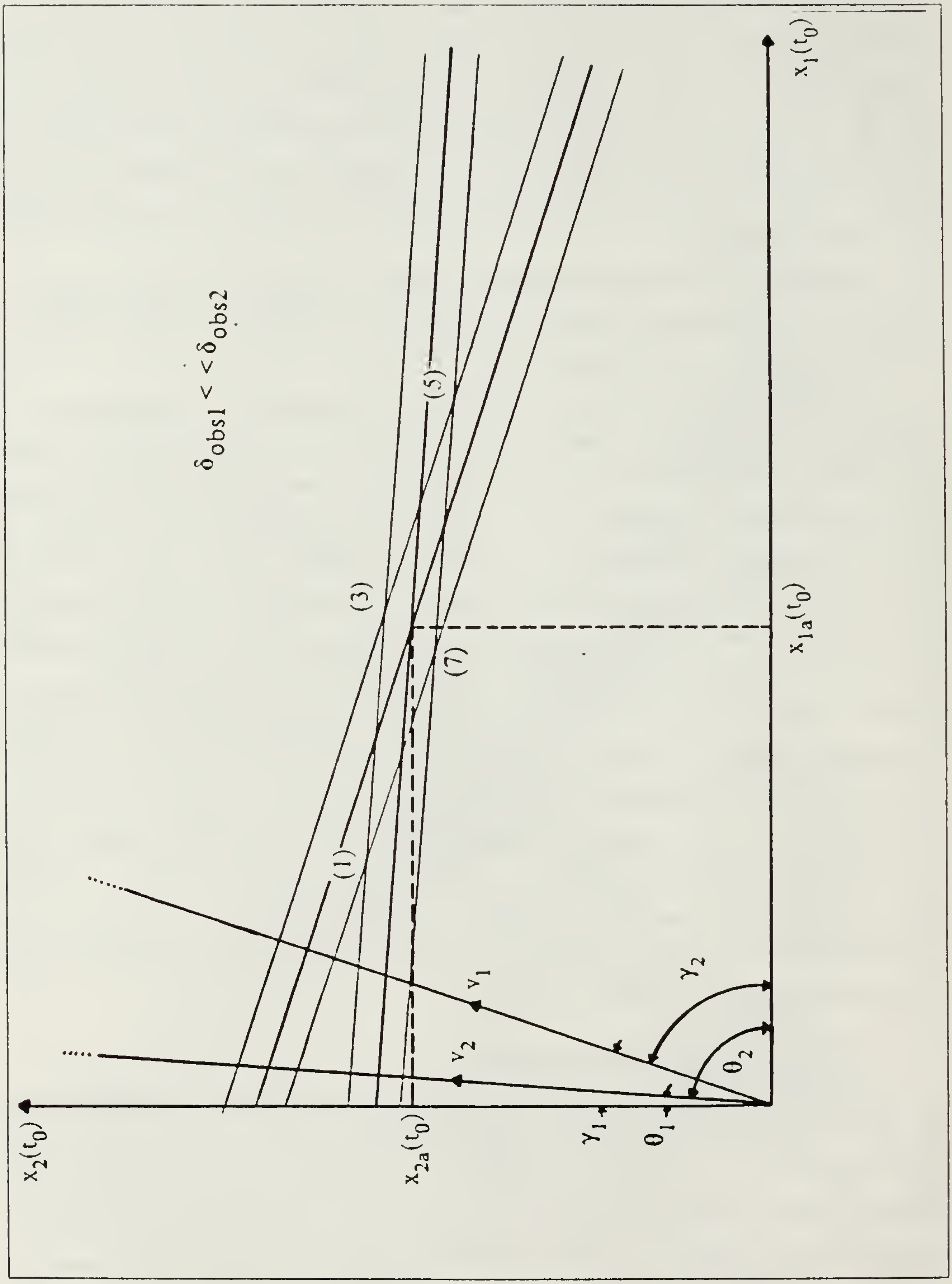

Figure 3.5 Importance of the angles on evaluating the degree of observability per state-variable.

spanned by Q^T one finds a way of comparing one particular vector against the state-space axes to obtain the relative degree of observability for each state variable ?

The direction of minimum error conveys considerable information about the relative observability of the system. Let $w_1, w_2, \ldots w_n$ be the columns of Q normalized. In [Ref. 7: p. 9] the author forms an observability function called ψ which is a scalar that is given by

$$\psi = (w_1{}^T u)^2 + (w_2{}^T u)^2 + \ldots + (w_n{}^T u)^2 \tag{3.26}$$

where u represents a vector in the direction of minimum error. The length of this vector is constrained to be of unit length, i.e.,

$$u^T u = 1 \tag{3.27}$$

Equation 3.26 can be written as follows :

$$\psi = u^T [w_1 w_1{}^T + w_2 w_2{}^T .. + w_n w_n{}^T] u \tag{3.28}$$

The quantity inside the brackets in equation 3.28 is actually $Q_n Q_n{}^T$ [Ref. 12: p. 220]. With this result in mind, this last equation can be written as

$$\psi = u^T (Q_n Q_n{}^T) u \tag{3.29}$$

The goal of this derivation is to find a vector that gives the direction of maximum error, yielding a minimum for the observability function ψ . To solve this problem, therefore, the gradient of ψ with respect to u must be found. In this problem there is a constraint that is given by equation 3.27. In general, any quadratic form of the type $x^T Px$ (P positive definite) is bounded below by the product of the minimum eigenvalue of P and the length of x and bounded above by the product of the maximum eigenvalue of P and the length of x [Ref. 13: p. 121]. For the particular case of equation 3.29, recalling that the length of u was assumed to be unity, this result can be written as

$$\lambda_{min} \leq u^T (Q_n Q_n{}^T) u \leq \lambda_{max} \tag{3.30}$$

where

λ_{min} is the smallest eigenvalue of $Q_nQ_n^T$

λ_{max} is the largest eigenvalue of $Q_nQ_n^T$

The same result can be obtained using direct differentiation of equation 3.29 with respect to u [Ref. 7: p. 10]. In this procedure the Lagrangian multiplier formulation is used where the constraint is given by equation 3.27. The Lagrangian multiplier formulation is given by equation 3.31.

$$\frac{d}{du}[u^T(Q_nQ_n^T)u - \lambda(u^Tu-1)] = 0 \tag{3.31}$$

The details on the differentiation of equation 3.31 can be found in [Ref. 14: p. 288]. The result of the differentiation is given by equation 3.32

$$(Q_nQ_n^T - \lambda I)u = 0 \tag{3.32}$$

Rearranging equation 3.32 and premultiplying both sides by u^T equation 3.33 is obtained.

$$u^T(Q_nQ_n^T)u = u^T\lambda u = \lambda \ = \psi \tag{3.33}$$

This equation shows that the minimum value of the observability function ψ is equal to the smallest eigenvalue of $Q_nQ_n^T$. Therefore, the direction of maximum error (where the observability function is minimum) will be given by the eigenvector associated with the smallest eigenvalue. The algorithm for the evaluation of the relative observability for each state variable is, now, evident and is as follows:

1. form Q_n , Q_n^T and multiply them together - call the product matrix QQT
2. find the eigenvalues and associated eigenvectors of QQT
3. find the smallest eigenvalue and associated eigenvector
4. sort the components of the eigenvector associated with the smallest eigenvalue in increasing order of magnitudes
5. associate the sorted components with the respective state variables
6. the smallest component is associated with the most observable state while the largest component is associated with the least observable state

G. SUMMARY

In this Chapter the concept of relative degree of observability was given as a measure of the accuracy involved in determining the initial states from the measurements of the output. The concept of an ill-conditioned matrix was presented as well as the problem involved in inverting highly ill-conditioned matrices. Based on the definition of relative degree of observability given and the description of the problem of an ill-conditioned matrix, four methods for evaluating the relative degree of observability were presented: the upper-bound error method, the standard deviation error method, the graphical method and the generalized graphical method. Examples were presented for the first two methods and algorithms were derived for the first, third and fourth methods. These algorithms are all imbedded in the same Fortran program (TOBS) presented in Appendix A.

IV. IMPROVEMENT OF THE DEGREE OF OBSERVABILITY OF A WEAKLY OBSERVABLE SYSTEM

A. INTRODUCTION

In the last Chapter, methods for evaluating the degree of observability for each state were presented. A natural extension of the study of this problem is the search for a method of improving the observability of a particular weakly observable state. In this Chapter methods for improvement will be presented, both for the single and multi-output cases.

B. THEORETICAL BACKGROUND

Before attempting to derive the method to increase the relative observability of a state variable, some important concepts of linear algebra need to be reviewed. Consider, then, the system represented by equation 2.1 where the square matrix A is such that all its eigenvalues are distinct. In this case the matrix A has at least one eigenvector for each distinct eigenvalue that satisfies equation 4.1

$$Ap=\lambda p \tag{4.1}$$

where $p\neq 0$ is the eigenvector of the square matrix A associated with the eigenvalue λ (λ is a scalar) [Ref. 10: p. 661]. Since the assumption is that the eigenvalues of A are all distinct, the n corresponding eigenvectors will all be linearly independent. Let $P=[p_1,\ldots,p_n]$ be the matrix of the eigenvectors of A. The matrix $M=P^{-1}$ exists because the columns of P are linearly independent. Let m_i^T be the rows of M. It can be shown that $A^Tm_i=\lambda_i m_i$ or equivalently $m_i^TA=\lambda_i m_i^T$. The m_i^T are often called the left (or row) eigenvectors of A in contrast to the right (or column) eigenvectors of A [Ref. 10: p. 664]. Note, however, that the so called left eigenvectors of A are nothing but the regular eigenvectors of A^T. In [Ref. 10: p. 664] the author shows two relations that will be quite important in the derivation of the method described in the next section. These relations are described as being the spectral decomposition of A and are given by

$$\begin{aligned} A &= \sum \lambda_i p_i m_i^T \\ &= P\Lambda M \\ &= M^{-1}\Lambda M \end{aligned} \tag{4.2}$$

$$\begin{aligned} A^2 &= \sum \lambda^2{}_i p_i m_i^T \\ &= P\Lambda^2 M \\ &= M^{-1}\Lambda^2 M \end{aligned} \tag{4.3}$$

In general the spectral decomposition of A^n is given by

$$\begin{aligned} A^n &= \sum \lambda^n{}_i p_i m_i^T \\ &= P\Lambda^n M \\ &= M^{-1}\Lambda^n M \end{aligned} \tag{4.4}$$

where $\Lambda =$ diag $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. The relations presented in the equations 4.2, 4.3 and 4.4 will be very helpful in understanding the method that is presented in the next section.

C. DERIVATION OF THE METHOD

Consider, now, the transpose of the observability matrix defined in equation 2.3 :

$$Q^T = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdots \\ CA^{n-1} \end{bmatrix} \tag{4.5}$$

Substituting the representation of the matrix A given by equations 4.2 and 4.3, yields equation 4.6 :

$$Q^T = \begin{bmatrix} CM^{-1}M \\ CM^{-1}\Lambda M \\ CM^{-1}\Lambda^2 M \\ \cdots \\ CM^{-1}\Lambda^{n-1}M \end{bmatrix} \tag{4.6}$$

Notice that in equation 4.6 the matrices C, M^{-1} and M are common to all the rows. Letting

$$L = CM^{-1}, \tag{4.7}$$

substituting equation 4.7 in equation 4.6 and post-factoring the matrix M, equation 4.8 is obtained :

$$Q^T = \begin{bmatrix} L \\ L\Lambda \\ L\Lambda^2 \\ \cdots \\ L\Lambda^{n-1} \end{bmatrix} M \tag{4.8}$$

The matrix L, is a row vector of dimension 1xn, because C is a 1xn row vector (in the case of single output) and M^{-1} is a nxn matrix. Thus, L can be written as

$$L = [L_1 \; L_2 \cdots L_n] \tag{4.9}$$

With this in mind and recalling that that Λ is a diagonal matrix whose elements are the eigenvalues of A, the matrix Q^T can be rewritten as follows :

$$Q^T = \begin{bmatrix} L_1 & L_2 & \cdots\cdots & L_n \\ L_1\lambda_1 & L_2\lambda_2 & \cdots\cdots & L_n\lambda_n \\ L_1\lambda_1{}^2 & L_2\lambda_2{}^2 & \cdots\cdots & L_n\lambda_n{}^2 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ L_1\lambda_1{}^{n-1} & L_2\lambda_2{}^{n-1} & \cdots\cdots & L_n\lambda_n{}^{n-1} \end{bmatrix} M \tag{4.10}$$

Looking carefully at equation 4.10 it can be seen that the matrix Q^T can be written as the product of three matrices :

$$Q^T = VGM \tag{4.11}$$

where V is the Vandermonde matrix of the eigenvalues of A:

$$V = \begin{bmatrix} 1 & 1 & \cdots\cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots\cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots\cdots & \lambda_n^2 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots\cdots & \lambda_n^{n-1} \end{bmatrix} \qquad (4.12)$$

G is a diagonal matrix whose elements are the entries of L:

$$G = \begin{bmatrix} L_1 & 0 & \cdots\cdots & 0 \\ 0 & L_2 & \cdots\cdots & 0 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ 0 & 0 & \cdots\cdots & L_n \end{bmatrix} \qquad (4.13)$$

and M is the matrix of the left eigenvectors of A.

With this procedure the matrix Q^T has been transformed into a product of three matrices. Since V and M depend only on the plant and G depends on both the plant and the C matrix, it seems to be G that needs to be studied and changed in order to improve the relative degree of observability of a particular state variable.

It was mentioned in Chapter III that the degree of ill-condition of Q^T is important in determining the magnitude of the errors (e_c) in computing the initial states once an error is introduced in measuring the output (e_m). Therefore, it is valuable to have a quantitative measure of the degree of ill-condition of the matrix Q^T. In [Ref. 15: p. 121] the author defines the condition number of a matrix W as the product of two matrix norms :

$$\mathrm{cond}(W) = ||W||\ ||W^{-1}|| \qquad (4.14)$$

If the condition number is small the matrix W is not ill-conditioned. If the matrix W is ill-conditioned the entries of its inverse will be large numbers, yielding a large condition number for W. However, this can also be true when the elements of W are small even in the absence of ill-conditioning. Multiplying the two norms has a normalization effect, so the condition number is large only for an ill conditioned system [Ref. 15: p.

122]. Observe that the condition number cannot be smaller than 1, which corresponds to the condition number of the identity matrix [Ref. 15: p. 121].

Equations 3.3 and 3.4 show that the vector of the measurements of the output consists of the sum of the actual values of the output $[y_a(t_0)]$ with the error (e_m) introduced while measuring the output. In other words $y_d(t_0) = y_a(t_0) + e_m$, which is equivalent to

$$e_m = y_d(t_0) - y_a(t_0) \tag{4.15}$$

But $y_a(t_0) = Q^T x_a(t_0)$ and $y_d(t_0) = Q^T x(t_0)$. By substituting these expressions in equation 4.15, equation 4.16 is obtained :

$$\begin{aligned} e_m &= Q^T x(t_0) - Q^T x_a(t_0) \\ &= Q^T [x(t_0) - x_a(t_0)] \\ &= Q^T e_c \end{aligned} \tag{4.16}$$

The error in the computed states is given by

$$e_c = (Q^T)^{-1} e_m \tag{4.17}$$

Taking norms and recalling that, in general, $||AB|| \leq ||A||\ ||B||$, yields for the particular case of equation 4.17 the result

$$||e_c|| \leq ||(Q^T)^{-1}||\ ||e_m|| \tag{4.18}$$

Applying norms to the last part of equation 4.16 the following is obtained :

$$||e_m|| \leq ||Q^T||\ ||e_c|| \tag{4.19}$$

Notice that equations 4.18 and 4.19 constitute, respectively, upper and lower bounds on the magnitude of the error in computing the initial states. This fact is shown in equation 4.20

$$\frac{||e_m||}{||Q^T||} \leq ||e_c|| \leq ||(Q^T)^{-1}||\ ||e_m|| \qquad (4.20)$$

Applying the same reasoning to the equations relating $x_a(t_0)$ to $y_a(t_0)$, yields for $x_a(t_0)$ the upper and lower bounds shown in equation 4.21,

$$\frac{||y_a(t_0)||}{||Q^T||} \leq ||x_a(t_0)|| \leq ||(Q^T)^{-1}||\ ||y_a(t_0)|| \qquad (4.21)$$

which is equivalent to the result obtained in equation 4.22

$$\frac{1}{||(Q^T)^{-1}||\ ||y_a(t_0)||} \leq \frac{1}{||x_a(t_0)||} \leq \frac{||Q^T||}{||y_a(t_0)||} \qquad (4.22)$$

Multiplying this last inequality by the inequality 4.20 and recalling the definition of condition number given in equation 4.14, equation 4.23 is obtained for the bounds of the relative error in computing the initial states.

$$\frac{1}{(\text{cond } Q^T)} \frac{||e_m||}{||y_a(t_0)||} \leq \frac{||e_c||}{||x_a(t_0)||} \leq (\text{cond } Q^T) \frac{||e_m||}{||y_a(t_0)||} \qquad (4.23)$$

This last result shows that the relative error in the computed solution (e_c) can be as large as the relative error in the measurements multiplied by the condition number. Of course it can also be as small as the relative error in the measurements divided by the condition number. So, when the condition number is large, the error in the measurements gives little information about the accuracy of $x(t_0)$. Conversely, if the condition number is nearly unity, the relative error in the measurements is a good measure of the relative error in computing the states [Ref. 15: p. 123], and the error in the measurements and the error in the calculated states are of the same order of magnitude. The goal of this whole procedure is to improve the relative degree of observability of a weakly observable state. Therefore, if the condition number of Q^T is minimized the relative error in the measurements will be of the same order of magnitude as the error in computing the states. The optimum result will be obtained for a condition number equal to one. In this case the upper and lower bounds on e_c will coincide. The condition number of Q^T is given by

$$\mathrm{cond}(Q^T) = ||Q^T|| \; ||(Q^T)^{-1}|| \tag{4.24}$$

Recalling equation 4.11 where $Q^T = VGM$ and substituting this in equation 4.24 yields

$$\begin{aligned} \mathrm{cond}(Q^T) &= ||VGM|| \; ||(VGM)^{-1}|| \\ &= ||VGM|| \; ||M^{-1}G^{-1}V^{-1}|| \end{aligned} \tag{4.25}$$

But the norm of the product of three matrices must not exceed the product of the three norms, i.e.,

$$\begin{aligned} \mathrm{cond}(Q^T) &\leq ||V|| \; ||G|| \; ||M|| \; ||M^{-1}|| \; ||G^{-1}|| \; ||V^{-1}|| \\ &\leq (\mathrm{cond}\ V)(\mathrm{cond}\ G)(\mathrm{cond}\ M) \end{aligned} \tag{4.26}$$

Observe that V, V^{-1}, M and M^{-1} cannot be changed because they are totally dependent on the dynamic characteristics of the plant. Therefore, only by minimizing the condition number of G will the condition number of Q^T be minimized. The matrix G is a function of the entries of C (equations 4.7 and 4.13), which are the only parts of the system that the user can change. On the other hand the condition number of Q^T is a function of the condition number of G (equation 4.26). Therefore, the relation that exists between the condition number of G and the condition number of Q^T can be seen as a function whose independent variable and abcissa is the condition number of G and the ordinate is the condition number of Q^T.

$$(\mathrm{cond}\ Q^T) = f(\mathrm{cond}\ G) \tag{4.27}$$

The goal of this procedure is to find the minimum of the condition number of Q^T by changing a specific entry in the C matrix. In other words, the goal is to find the minimum of f. In order to obtain an algorithm that accomplishes this, the function f must be studied in more detail. Both the abcissa and the ordinate are condition numbers of matrices which, by definition, can never be less than one. Only in very special cases is the condition number of a matrix equal to one. The identity matrix and all its multiples are among those special cases. Notice, also that the first row of Q^T is the C matrix in the single output case. Therefore, only when c_{11} is different from zero and all the other entries in the C matrix are zero, will the possibility of obtaining a minimum for the condition number of Q^T equal to one be likely to occur. In other words, in the majority of cases both abcissa and ordinate will be greater than one. The matrix G plays an important role in the development of the algorithm that minimizes the condition number of Q^T by changing a specific entry in the C matrix. In order to understand how the matrix G is used in this derivation a more detailed study of its entries will be done. The matrix G is a diagonal matrix whose elements are given by

$$\begin{aligned} L_1 &= g_{11} = \sum c_{1j} m_{j1} \\ L_2 &= g_{22} = \sum c_{1j} m_{j2} \\ &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ L_s &= g_{ss} = \sum c_{1j} m_{js} \\ &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \qquad\qquad j = 1,2,\ldots,n \qquad\qquad (4.28) \\ &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ L_\ell &= g_{\ell\ell} = \sum c_{1j} m_{j\ell} \\ &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ L_n &= g_{nn} = \sum c_{1j} m_{jn} \end{aligned}$$

where the c_{1j} are the entries of the C matrix and the m_{ji} are the entries of the M^{-1} matrix.

The norm used to compute the condition number may be any matrix norm [Ref. 15: p.118]. In the present case the maximum column sum norm will be used to compute the condition number and will be defined as follows :

$$||G||_1 = \max \sum |g_{jj}| \qquad j = 1,2,\ldots,n \qquad (4.29)$$

Notice that both subscripts of g are equal because the matrix G is diagonal. In this particular case the sum is equal to the only non-zero entry in the column. The norm is then given by equation 4.30:

$$||G||_1 = \max |g_{jj}| \qquad j = 1,2,\ldots,n \qquad (4.30)$$

Suppose that the maximum element in G is $g_{\ell\ell}$ where $g_{\ell\ell}$ is given by equation 4.28. According to equation 4.30 the norm of G will then be equal to

$$||G||_1 = |g_{\ell\ell}| \qquad (4.31)$$

Since the G matrix is diagonal its minimum entry corresponds to the maximum entry in G^{-1} which is also diagonal. Its norm is given by equation 4.32 :

$$||G^{-1}|| = \frac{1}{\min|g_{jj}|} \qquad j = 1,2,\ldots,n \tag{4.32}$$

or

$$||G^{-1}|| = \frac{1}{|g_{ss}|} \tag{4.33}$$

where g_{ss} is the smallest entry in G and is given by equation 4.28 . The condition number of G is, by definition, given by the product of the two norms given in equations 4.31 and 4.33. Substituting equation 4.28 for the values of g_{ss} and $g_{\ell\ell}$ in equations 4.31 and 4.33 yields, for the equation of the condition number of G

$$\frac{\sum_{j=1}^{n} c_{1j} m_{j\ell}}{\sum_{j=1}^{n} c_{1j} m_{js}} = (\text{cond } G) \tag{4.34}$$

The ratio given by equation 4.34 is also the ratio between the largest and smallest eigenvalues of G (G is a diagonal matrix), which is another definition of condition number of a matrix [Ref. 16: p. 38]. The right-hand side of equation 4.34 is a constant. Let the inverse of that constant be equal to K. Equation 4.34 can then be rewritten as follows:

$$K \sum_{j=1}^{n} c_{1j} m_{j\ell} = \sum_{j=1}^{n} c_{1j} m_{js} \tag{4.35}$$

Since there is only one equation (equation 4.35) to provide the changes in the C matrix to improve the relative degree of observability of the system, a decision has to be made upon the entry in the C matrix for which equation 4.35 will be solved. Assuming that at this stage one has already computed the relative degree of observability of the system, one knows which is the least observable state. If that state is state x_r, equation 4.35 will be solved for c_{1r} because this is the entry in the C matrix associated with the least observable state. Expanding equation 4.35 yields :

$$K(c_{11}m_{1\ell}+\ldots+c_{1r}m_{r\ell}+\ldots+c_{1n}m_{n\ell})= \\ c_{11}m_{1s}+\ldots+c_{1r}m_{rs}+\ldots+c_{1n}m_{ns} \tag{4.36}$$

Rearranging and collecting common terms,

$$(Km_{1\ell}-m_{1s})c_{11}+\ldots+(Km_{r\ell}-m_{rs})c_{1r}+\ldots+(Km_{n\ell}-m_{ns})c_{1n}=0 \tag{4.37}$$

Solving for c_{1r} equation 4.38 is obtained

$$c_{1r} = -\frac{1}{\beta_r}\sum_{\substack{i=1\\i\neq r}}^{n}\beta_i c_{1i}, \qquad \beta_r\neq 0 \tag{4.38}$$

where in general $\beta_i=(Km_{i\ell}-m_{is})$. At this point every tool is available to develop the algorithm that will provide the new entry in the C matrix corresponding to the minimum condition number of Q^T.

The initialization of the algorithm is done by selecting an abcissa equal to the condition number of G of the initial system and an ordinate equal to the condition number of Q^T of the same system. To define the interval where the minimum is to be evaluated, another point has to be chosen. As previously mentioned, the condition number of a matrix is a quantity that is always greater or equal to one. With this in mind, the choice of the second point selected for the initialization of the algorithm is the one where the abcissa (condition number of G) is equal to one and the ordinate is the corresponding value for the condition number of Q^T. The computation of the condition number of Q^T given the condition number of G is a three step procedure that can be summarized as follows:

STEP-1 : Given the condition number of G (abcissa) solve equation 4.38 for c_{1r}.

STEP-2 : Find the new Q^T by using equation 2.3 and transposing it.

STEP-3 : Find the condition number of the new Q^T (new ordinate).

A flow-chart of the algorithm with the initialization and minimization procedure is presented in Figure 4.1. In the flow-chart the letter f corresponds to the definition in equation 4.27 and represents the three step procedure described above. X represents the condition number of G while Y represents the condition number of Q^T. After bisecting the interval a decision has to be made concerning where to search for the minimum of f, i.e., if to the right or to the left of X3. The information given by the derivative of f near X3,Y3 is used to make this decision: if the derivative is positive the interval to be searched is to the left of X3. If the derivative is negative the interval to be searched is to the right of X3. Once the direction of search is known, the points limiting the search interval are renamed: the point with the largest ordinate is X1,Y1 and the point with the smallest ordinate is X2,Y2. If the stopping criterion is met the algorithm stops. Otherwise, the bisecting procedure is executed again. This algorithm gives a local minimum. Other initializations have to be used to try to find other minimums. The whole algorithm containing the required minimization of the condition number of Q^T, was implemented in a Fortran computer program called COBS and is presented in the Appendix B.

The choice of the entry of the C matrix associated with the least observable state as the one to be changed is made here. However, this issue is controversial because real cases may happen where the user cannot change the C matrix according to the least observable state. Furthermore, an independent change in an entry of the C matrix may not be possible, i.e., a change in one entry of C may only be possible if another entry is changed. Several combinations are possible. Therefore, according to a specific case, the user may have to try several allowable changes in the C matrix to determine the minimum condition number of Q^T. The program that is in Appendix B has the default feature of changing the entry of the C matrix associated with the least observable state. However, any other entry in the C matrix can be changed in order to increase the degree of observability of a specific state variable. This is accomplished by following the documentation contained in the listing of COBS.

Once the condition number of Q^T has been minimized and the new entries of the C matrix found, the system's observability has to be checked. In equation 4.11 it was seen that Q^T could be decomposed into the product of three matrices. The Vandermonde and the matrix of the left eigenvectors are full rank matrices because by assumption the plant has distinct eigenvalues. Therefore, if G is a full rank matrix Q^T will also be a full rank matrix and the system will be completely observable. For G to be full rank all its entries must be different from zero because G is a diagonal matrix.

D. THE MULTI-OUTPUT CASE

What has been said so far is only valid for the single output case. Since this is very seldom the case, the multi-output situation will be considered in this section.

Consider, again, the linear time invariant system described by equations 2.1 and 2.2 where C is a mxn matrix and $m \geq 1$. The observability matrix Q is, for the general case of multi-output, given by

$$Q = [C^T \quad A^T C^T \quad (A^T)^2 C^T \quad \dots \quad (A^T)^{n-\ell} C^T] \tag{4.39}$$

where ℓ is the rank of C [Ref. 17: p. 503]. This new observability matrix is nxp, where $p=(n-\ell+1)m$. Notice that if $m=1$ and $\ell=1$ (single output case), the matrix Q will be square. The fact that the matrix Q is no longer square, transforms a relatively simple problem into a much more elaborate one. Recalling equation 3.2 where the relation between the derivatives of the output and the states is given, if one attempts to solve that system of equations for $x(t_0)$ when Q is not a square matrix, the concept of generalized or pseudo-inverse must be introduced. Premultiplying equation 3.2 by Q yields

$$Qy_d(t_0) = QQ^T x(t_0) \tag{4.40}$$

Solving equation 4.40 for $x(t_0)$.

$$x(t_0) = (QQ^T)^{-1} Q y_d(t_0) \tag{4.41}$$

The matrix $(QQ^T)^{-1}Q$ is called the pseudo-inverse of Q^T [Ref. 17: p. 577]. The solution of equation 3.2 using the pseudo-inverse is an approximate solution in the least-squares sense [Ref. 17: p. 578]. Let $QQ^T = R$. The problem is, now, finding how invertible the the R matrix is, i.e., how ill-conditioned the matrix R is. One possible procedure is the one used for the single output case. Thus, performing the spectral factorization of the matrix A in the transpose of equation 4.39 by substituting its representation given by equations 4.2, 4.3 and 4.4 yields equation 4.42:

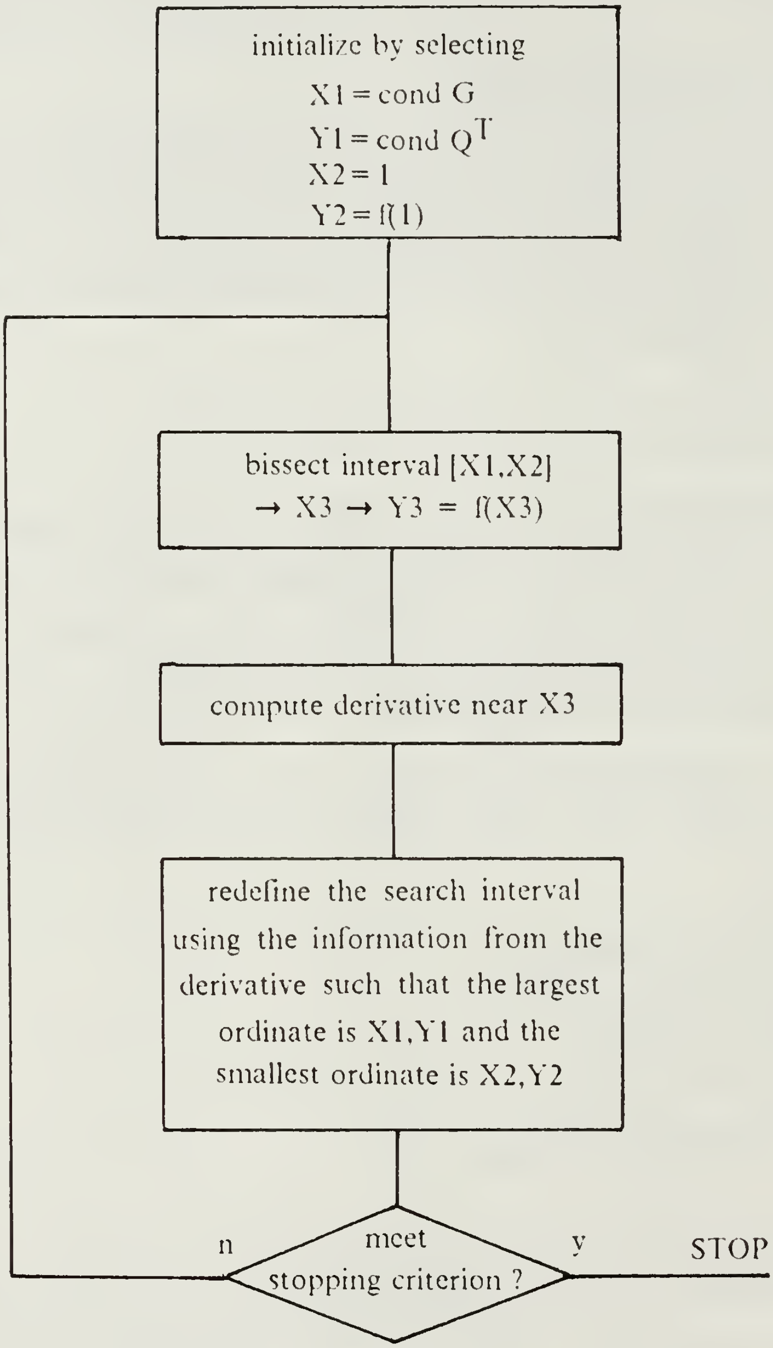
initialize by selecting
X1 = cond G
Y1 = cond Q^T
X2 = 1
Y2 = f(1)
bissect interval [X1,X2]
→ X3 → Y3 = f(X3)
compute derivative near X3
redefine the search interval using the information from the derivative such that the largest ordinate is X1,Y1 and the smallest ordinate is X2,Y2
meet stopping criterion ?
n
y
STOP

Figure 4.1 Algorithm to provide the minimum of the condition number of Q^T by changing a specific entry in C.

$$Q^T = \begin{bmatrix} CM^{-1}M \\ CM^{-1}AM \\ CM^{-1}A^2M \\ \cdots \\ CM^{-1}A^{n-\ell}M \end{bmatrix} \tag{4.42}$$

Notice, again, that as in the single output case, the matrices C, M^{-1} and M are common to all the rows. Let

$$L = CM^{-1} \tag{4.43}$$

where the dimension of L is mxn and

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots\cdots & c_{1n} \\ c_{21} & c_{22} & \cdots\cdots & c_{2n} \\ c_{31} & c_{32} & \cdots\cdots & c_{3n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ c_{m1} & c_{m2} & \cdots\cdots & c_{mn} \end{bmatrix} \tag{4.44}$$

and

$$M^{-1} = \begin{bmatrix} m_{11} & m_{12} & \cdots\cdots & m_{1n} \\ m_{21} & m_{22} & \cdots\cdots & m_{2n} \\ m_{31} & m_{32} & \cdots\cdots & m_{3n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ m_{n1} & m_{n2} & \cdots\cdots & m_{nn} \end{bmatrix} \tag{4.45}$$

Substituting equation 4.43 in equation 4.42 and post-factoring the matrix M, equation 4.46 is obtained:

$$Q^T = \begin{bmatrix} L \\ L\Lambda \\ L\Lambda^2 \\ \cdots \\ L\Lambda^{n-\ell} \end{bmatrix} M \qquad (4.46)$$

Performing all the matrix multiplications in equation 4.46 except those involving the matrix M, the following representation for Q^T is obtained:

$$Q^T = \left[\begin{array}{cccc} \zeta_{11} & \zeta_{12} & \cdots\cdots & \zeta_{1n} \\ \zeta_{21} & \zeta_{22} & \cdots\cdots & \zeta_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \zeta_{m1} & \zeta_{m2} & \cdots\cdots & \zeta_{mn} \\ \hline \zeta_{11}\lambda_1 & \zeta_{12}\lambda_2 & \cdots\cdots & \zeta_{1n}\lambda_n \\ \zeta_{21}\lambda_1 & \zeta_{22}\lambda_2 & \cdots\cdots & \zeta_{2n}\lambda_n \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \zeta_{m1}\lambda_1 & \zeta_{m2}\lambda_2 & \cdots\cdots & \zeta_{mn}\lambda_n \\ \hline \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \hline \zeta_{11}\lambda_1^{n-\ell} & \zeta_{12}\lambda_2^{n-\ell} & \cdots\cdots & \zeta_{1n}\lambda_n^{n-\ell} \\ \zeta_{21}\lambda_1^{n-\ell} & \zeta_{22}\lambda_2^{n-\ell} & \cdots\cdots & \zeta_{2n}\lambda_n^{n-\ell} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \zeta_{m1}\lambda_1^{n-\ell} & \zeta_{m2}\lambda_2^{n-\ell} & \cdots\cdots & \zeta_{mn}\lambda_n^{n-\ell} \end{array}\right] M \qquad (4.47)$$

where ζ_{ij} are the entries of L and λ_i, the system eigenvalues, are the entries of the diagonal of Λ. The goal is to decompose the first matrix on the right hand side of equation 4.47 as the product of two matrices such that only one of them depends on the C matrix. The matrix Q^T can actually be written as the product of three matrices:

$$Q^T = \underline{V}\,\underline{G}M \qquad (4.48)$$

where $\underline{V}$ has a Vandermonde structure,

$$\underline{V} = \begin{bmatrix} I & I & \cdots\cdots & I \\ \hdashline \Lambda_1 & \Lambda_2 & \cdots\cdots & \Lambda_n \\ \hdashline \Lambda_1{}^2 & \Lambda_2{}^2 & \cdots\cdots & \Lambda_n{}^2 \\ \hdashline \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \hdashline \Lambda_1{}^{n-\ell} & \Lambda_2{}^{n-\ell} & \cdots\cdots & \Lambda_n{}^{n-\ell} \end{bmatrix} \qquad (4.49)$$

Each matrix forming $\underline{V}$ is a mxm matrix of the form

$$\Lambda_j^{i-1} = \begin{bmatrix} \lambda_j^{i-1} & 0 & \cdots\cdots & 0 \\ 0 & \lambda_j^{i-1} & \cdots\cdots & 0 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ 0 & 0 & \cdots\cdots & \lambda_j^{i-1} \end{bmatrix} \qquad (4.50)$$

where i, the $\underline{V}$ super-row index, is equal to 2,3,...,n and j, the $\underline{V}$ super-column index is equal to 1,2,...,n. Therefore, $\underline{V}$ has in each of its super-columns (n-ℓ+1) matrices with m rows and m columns. On the other hand, there are n submatrices in each super-row of $\underline{V}$. Thus, the matrix $\underline{V}$ has p rows and q columns where $p=(n-\ell+1)m$ and $q=mn$. The matrix I is an mxm identity matrix. The matrix $\underline{G}$ is a qxn block diagonal matrix

$$\underline{G} = \begin{bmatrix} \underline{L}_1 & 0 & \cdots\cdots & 0 \\ \hdashline 0 & \underline{L}_2 & \cdots\cdots & 0 \\ \hdashline \vdots & \vdots & \vdots & \vdots \\ \hdashline 0 & 0 & \cdots\cdots & \underline{L}_n \end{bmatrix} \qquad (4.51)$$

Each matrix contained in $\underline{G}$ is actually an mx1 column vector corresponding to the columns of the matrix L given by equation 4.43. In general, each submatrix of $\underline{G}$ is given by

$$\underline{L}_i = \begin{bmatrix} \zeta_{1i} \\ \zeta_{2i} \\ \zeta_{3i} \\ \cdots \\ \zeta_{mi} \end{bmatrix} \qquad i=1,2,\ldots,n \qquad (4.52)$$

This decomposition of the Q^T matrix seems to be complicated and even difficult to use. Two examples will be presented to clarify the ideas. Suppose that a third order plant is given in which only one output is available, i.e., $n=3$, $m=1$ and the rank of C is $\ell=1$. The dimension of Q^T is pxn where $p=(n-\ell+1)m=3$ and $n=3$; $\underline{V}$ is a pxq matrix where $p=3$ and $q=nxm=3$; $\underline{G}$ is qxn where $q=3$ and $p=3$. In summary Q^T, $\underline{V}$ and $\underline{G}$ are 3x3 matrices. The matrix $\underline{V}$ has in each of its columns $n-\ell+1=3$ matrices with $m=1$ row each. According to equation 4.49 the first row matrices are identity matrices which have m rows (in this case one row), i.e., the first row elements are three ones. The second row elements are found by using equation 4.50 along with equation 4.49. Since one is looking for the second super-row of $\underline{V}$, the index i is equal to 2 and the super-column index j is equal to 1,2 and 3. Therefore, the matrices that constitute the second super-row of $\underline{V}$ have $m=1$ row and have entries λ_1, λ_2 and λ_3. The third and last super-row matrices of $\underline{V}$ are λ_1^2, λ_2^2 and λ_3^2.

$$\underline{V} = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix} \qquad (4.53)$$

The matrix $\underline{G}$ is a qxn, i.e., 3x3 diagonal block matrix whose entries are given by equation 4.52. Since $m=1$ the $\underline{L}_i$ are scalars and are equal to $\underline{L}_1=\zeta_{11}$, $\underline{L}_2=\zeta_{12}$, $\underline{L}_3=\zeta_{13}$. The matrix $\underline{G}$ will be equal to

$$\underline{G} = \begin{bmatrix} \zeta_{11} & 0 & 0 \\ 0 & \zeta_{12} & 0 \\ 0 & 0 & \zeta_{13} \end{bmatrix} \qquad (4.54)$$

As can be seen by equations 4.53 and 4.54, by applying the spectral decomposition of Q^T for this generalized case, a result is obtained that agrees with the one that would be obtained if equations 4.12 and 4.13 were used.

Suppose, now that the system is third order with two outputs. In other words, $n=3$ and $m=2$. The rank of C is the smaller of m and n, i.e., $\ell=2$. Thus Q^T will be a $(n-\ell+1)m\text{x}n$ matrix, i.e., Q^T will be a 4x3 matrix. This matrix will be decomposed in the product of three matrices $\underline{V}$, $\underline{G}$ and M whose dimensions are as follows: $\underline{V}$ is a pxq matrix where $p=(n-\ell+1)m=4$ and $q=m\text{x}n=6$. The matrix $\underline{G}$ is $q\text{x}n=6\text{x}3$ while the matrix M is $n\text{x}n=3\text{x}3$. Each submatrix of $\underline{V}$ is $m\text{x}m=2\text{x}2$. Therefore, $\underline{V}$ has two super rows and three super columns because that is the only way of obtaining a dimension of 4x6 for the matrix $\underline{V}$. The submatrices in the first super-rows of $\underline{V}$ are 2x2 identity matrices. Therefore, in the first super-row of $\underline{V}$ there are three (n) 2x2 (mxm) identity matrices. The second super-row is computed using equation 4.50 with i (the super-row index) equal to two and j (the super-column index) running from one to three since there are three super columns in the $\underline{V}$ matrix.

$$\underline{V} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ \lambda_1 & 0 & \lambda_2 & 0 & \lambda_3 & 0 \\ 0 & \lambda_1 & 0 & \lambda_2 & 0 & \lambda_3 \end{bmatrix} \tag{4.55}$$

The entries of matrix $\underline{G}$ are given by equation 4.52. Since $m=2$ the $\underline{L}_i$ are 2x1 column vectors. The matrix $\underline{G}$ will have three (n) of these column vectors in its diagonal.

$$\underline{G} = \begin{bmatrix} \zeta_{11} & 0 & 0 \\ \zeta_{21} & 0 & 0 \\ 0 & \zeta_{12} & 0 \\ 0 & \zeta_{22} & 0 \\ 0 & 0 & \zeta_{13} \\ 0 & 0 & \zeta_{23} \end{bmatrix} \tag{4.56}$$

If one multiplies equation 4.55 ($\underline{V}$) by equation 4.56 ($\underline{G}$) the result obtained will be equivalent to the one that would be obtained if equation 4.47 is used for the same values of n, m and ℓ.

In order to find the initial states the inversion of a matrix must be performed, as was mentioned at the beginning of this Section. The matrix $R=QQ^T$ is the one that has to be inverted as shown in equation 4.41. Thus, for the multiple-output case the condition number of R is the value that gives a measure of how observable the state variables are. Expanding R the following is obtained:

$$\begin{aligned} R &= QQ^T \\ &= (\underline{V}\underline{G}M)^T(\underline{V}\underline{G}M) \\ &= M^T\underline{G}^T\underline{V}^T\underline{V}\underline{G}M \end{aligned} \tag{4.57}$$

In general, any quadratic form of the type x^TPx (P positive definite) is bounded below by the product of the minimum eigenvalue of P and the length of x and bounded above by the product of the maximum eigenvalue of P and the length of x [Ref. 13: p. 121]. For the particular case of equation 4.57 this result can be written as

$$\lambda_{min}||x||^2 \le x^TQQ^Tx \le \lambda_{max}||x||^2 \tag{4.58}$$

where λ_{max} and λ_{min} are the maximum and minimum eigenvalues of the square matrix $R = QQ^T$. The theorem mentioned in [Ref. 13: p. 121] can be used several times in equation 4.58 until upper and lower bounds are obtained only as functions of the maximum and minimum eigenvalues of the square matrices $\underline{V}^T\underline{V}$, $\underline{G}^T\underline{G}$ and M^TM. Equation 4.59 presents what happens for the upper bound

$$\begin{aligned} xQQ^Tx &= x^TM^T\underline{G}^T(\underline{V}^T\underline{V})\underline{G}Mx \\ &\le \lambda_{max}(\underline{V}^T\underline{V})||x^TM^T(\underline{G}^T\underline{G})Mx|| \\ &\le \lambda_{max}(\underline{V}^T\underline{V})\lambda_{max}(\underline{G}^T\underline{G})||x^TM^TMx|| \\ &\le \lambda_{max}(\underline{V}^T\underline{V})\lambda_{max}(\underline{G}^T\underline{G})\lambda_{max}(M^TM)||x||^2 \end{aligned} \tag{4.59}$$

and an equivalent result can be obtained for the lower bound, i.e.,

$$xQQ^Tx \ge \lambda_{min}(\underline{V}^T\underline{V})\lambda_{min}(\underline{G}^T\underline{G})\lambda_{min}(M^TM)||x||^2 \tag{4.60}$$

Recalling the fact that in general the norm of the product is less or equal than the product of the norms, the three results presented in equations 4.58, 4.59 and 4.60 imply that

$$\lambda_{max}(QQ^T) \le \lambda_{max}(\underline{V}^T\underline{V})\lambda_{max}(\underline{G}^T\underline{G})\lambda_{max}(M^TM) \tag{4.61}$$

and

$$\lambda_{\min}(QQ^T) \geq \lambda_{\min}(\underline{V}^T\underline{V})\lambda_{\min}(\underline{G}^T\underline{G})\lambda_{\min}(M^TM) \qquad (4.62)$$

In [Ref. 16: p. 38] the author mentions that the ratio between the largest and smallest eigenvalues is a measure of how ill-conditioned a matrix is. With this in mind and looking at the inequalities 4.61 and 4.62 it can be said that

$$\text{cond}(QQ^T) \leq [\text{cond}(\underline{V}^T\underline{V})][\text{cond}(\underline{G}^T\underline{G})][\text{cond}(M^TM)] \qquad (4.63)$$

Observe that although the condition number of QQ^T is a function of $\underline{V}^T\underline{V}$, $\underline{G}^T\underline{G}$ and M^TM., any improvement of the degree of observability of the system can only be accomplished by changing the condition number of $\underline{G}^T\underline{G}$ because this is the only part of the product of the three matrices that is not totally dependent on the plant. Assuming that the user is only allowed to alter the C matrix and recalling equation 4.43, it can be seen that only by minimizing the condition number of $\underline{G}^T\underline{G}$ will the condition number of QQ^T be minimized. The matrix $\underline{G}^T\underline{G}$ is a nxn diagonal matrix whose entries are actually the lengths of the column vectors of $\underline{G}$ squared. Let the entries of $\underline{G}^T\underline{G}$ be the g_{jj}

$$\begin{array}{l} g_{11} = \sum \zeta_{i1}^2 \\ g_{22} = \sum \zeta_{i2}^2 \\ \cdot \quad \cdot \quad \cdot \\ g_{nn} = \sum \zeta_{in}^2 \end{array} \qquad i = 1,2,\ldots,m \qquad (4.64)$$

Thus, the condition number of $\underline{G}^T\underline{G}$ as it is defined in [Ref. 16: p. 38], will be given by the ratio of the largest and the smallest entries in $\underline{G}^T\underline{G}$. The condition number of $\underline{G}^T\underline{G}$ is given by equation 4.65:

$$\frac{\sum\limits_{i=1}^{m} \zeta_{i\ell}^{\,2}}{\sum\limits_{i=1}^{m} \zeta_{is}^{\,2}} = (\text{cond } \underline{G}^T\underline{G}) \qquad (4.65)$$

where $\zeta_{i\ell}$ is the largest element in the $\underline{G}^T\underline{G}$ matrix while ζ_{is} is the smallest element in the same matrix. Comparing equation 4.65 with equation 4.34 and recalling the fact that the numerator and denominator of equation 4.34 are respectively the largest and smallest entries in the matrix G (which is the equivalent to the matrix $\underline{G}$ for the multi-output case), it can be seen that starting from two different definitions of condition number, two equivalent equations for the same quantity are obtained. The entries of the matrix $\underline{G}$ are also summations and are given , in general, by

$$\zeta_{ij} = \sum_{p=1}^{n} c_{ip} m_{pj} \quad i = 1,2,\ldots,m \quad j = 1,2,\ldots,n \tag{4.66}$$

Substituting equation 4.66 in equation 4.65 for the values of $\zeta_{i\ell}$ and ζ_{is} yields

$$\frac{\sum_{i=1}^{m} \left(\sum_{p=1}^{n} c_{ip} m_{p\ell}\right)^2}{\sum_{i=1}^{m} \left(\sum_{p=1}^{n} c_{ip} m_{ps}\right)^2} = (\text{cond } \underline{G}^T\underline{G}) \tag{4.67}$$

At this point, if one attempts to use the same method used in the single output case, namely solving equation 4.67 for the entry of C that is related to the state whose relative observability is to be improved, it can be seen that this task is far from being an easy one because the equations that have to be solved are quadratics. Therefore, one of the possible solutions for this problem is the following: the first assumption is that the user, after studying the possibilities of the system under test, knows already which entries of C can be changed and how much can they be changed. Then only one of the entries of C would be changed in each run over the allowed interval. This procedure would have to be repeated for every entry in the C matrix that can be changed. Meanwhile, using equation 4.67 the condition number of $\underline{G}^T\underline{G}$ would be computed as well as the condition number of QQ^T. For every change in the entry of C under test the relative degree of observability of the system would be computed. A plot of the condition number of $\underline{G}^T\underline{G}$ against the condition number of QQ^T for every possible allowable change in the C matrix would be an extremely valuable output. With this information and the knowledge of what states are more observable under different conditions, a choice of the best allowable combination of the entries of C could then be made such that the condition number of QQ^T is minimum or a specific state variable is more observable than the others. Notice the parallelism that exists between the derivations involving the multi-output case and the generalized graphical method developed in Chapter III to evaluate the relative degree of observability of the state-variables. In both cases the matrix QQ^T was involved as the center of the problem: a search for the minimum eigenvector in the method of Chapter III and the search for the minimum condition number of QQ^T in the method presented in this Chapter. The goals of both methods are somehow different although correlated because the second uses information from the first one.

E. CASE WHERE THERE IS A CONTROL INPUT

In the beginning of Chapter III the concept of relative observability was defined as a measure of how accurately one can determine the initial states from the measurements of the output at time $t = t_1, t_2, \ldots, t_n$. The control input u(t) was set to zero because the effect of the input signal can always be determined if $x(t_0)$ can be found. Now, considering the existence of a control input $u(t) \neq 0$, the concept of relative observability must be redefined as follows: measuring y(t) and u(t) at time $t = t_1, t_2, \ldots, t_n$ one wants to be able to evaluate which state components are most and least accurately determined from the measurements of both the input and output. It will be shown in this section that the fact that one is measuring the input signal as well as the output signal doesn't alter in any way what has been shown to be valid for the case of $u(t) = 0$. Consider, again, the LTI system described by equations 2.1 and 2.2. Differentiating equation 2.2 n-1 times and substituting x for its value given in equation 2.1, the following is obtained:

$$\begin{aligned}
y(t_0) &= Cx(t_0) \qquad (4.68)\\
\dot{y}(t_0) &= CAx(t_0) + CBu(t_0)\\
\ddot{y}(t_0) &= CA^2x(t_0) + CABu(t_0) + CB\dot{u}(t_0)\\
\dddot{y}(t_0) &= CA^3x(t_0) + CA^2Bu(t_0) + CAB\dot{u}(t_0) + CB\ddot{u}(t_0)\\
&\vdots \qquad \vdots \qquad \vdots\\
\overset{(n-1)}{y}(t_0) &= CA^{n-1}x(t_0) + CA^{n-2}Bu(t_0) + CA^{n-3}Bu(t_0) + \ldots + CB\overset{(n-2)}{u}(t_0)
\end{aligned}$$

Notice that the arguments of the input and its derivatives is t_0 which means that the derivatives of the inputs are being evaluated at time $t = t_0$. Equation 4.68 can be written in the matrix form

$$y_d(t_0) = Q^T x(t_0) + \Omega u_d(t_0) \tag{4.69}$$

where:

$y_d(t_0)$ is a nx1 column vector of the output and its n-1 first derivatives,

$x(t_0)$ is a nx1 column vector of the state variables at time t_0

Ω is a nx(n-1) lower triangular Toeplitz matrix that represents the impulse response of the system and is given by equation 4.70

$$\Omega = \begin{bmatrix} 0 & 0 & \dots & 0 \\ CB & 0 & \dots & 0 \\ CAB & CB & \dots & 0 \\ \dots & \dots & \dots & \dots \\ CA^{n-2}B & CA^{n-3}B & \dots & CB \end{bmatrix} \tag{4.70}$$

and $u_d(t_0)$ is a 1x(n-1) column vector of the input and its n-2 first derivatives. Considering, now, the single output case, equation 4.69 is solved for $x(t_0)$

$$x(t_0) = (Q^T)^{-1} y_d(t_0) - (Q^T)^{-1} \Omega u_d(t_0) \tag{4.71}$$

Observe that the projection of an error made in measurements of y(t) and u(t) in computing the initial states is determined, again, by the degree of ill-conditioning of Q^T which is the only matrix that needs to be inverted in this process. The matrix **Ω** causes no problem to because it can be determined very accurately (assuming that the identification of the plant has been done). Consider the example used in Chapter III to explain the graphical method where

$$Q^T = \begin{bmatrix} 1.01 & 0.99 \\ 0.99 & 1.01 \end{bmatrix}$$

The matrix **Ω** was equal to zero because the control input was set to zero. Now the control input is different from zero which implies that **Ω** is also different from zero. Let **Ω** be equal to

$$\Omega = \begin{bmatrix} 0 \\ \alpha \end{bmatrix}$$

where alpha is any real number different from zero. Substituting the matrices Q^T and Ω in equation 4.69 yields

$$1.01x_1(t_0) + 0.99x_2(t_0) = y_1(t_0) \tag{4.72}$$

$$0.99x_1(t_0) + 1.01x_2(t_0) = y_2(t_0) - \alpha u_d(t_0) \tag{4.73}$$

Equations 4.72 and 4.73 can be plotted in the $x_1(t_0)$ $x_2(t_0)$ plane as was done in Chapter III Section E. However, just by inspection one can see that the slopes of the lines corresponding to equations 4.72 and 4.73 are the same as the slopes of the lines corresponding to equations 3.20 and 3.21 although a new member exists in equation 4.73. This new member appears due to the existence of a control input and is only responsible for a change in the intercepts, i.e., for a change in the solution of the system of equations of the example in Chapter III. Since the slopes are equal so are the angles between the lines. It was shown in Chapter III that the angles between these lines are highly correlated with the degree of ill-conditioning of the Q^T matrix: the smaller the angles the higher the degree of ill-conditioning of the matrix Q^T. Therefore, the fact that a control input exists doesn't bring any changes in the basic ideas presented so far leading to the evaluation of the relative degree of observability of the state variables of a given system. The same thing happens for the improvement of the degree of observability of a weakly observable state. In the case of $u(t)=0$ the procedure was the factorization of the Q^T matrix as the product of three full rank matrices such that only one was dependent on the C matrix, which was the only part of the system that the user is allowed to change. Then, the goal was to find a new entry in the C matrix such that the condition number of Q^T was minimum. Once a minimum was found the relative observability tests were performed as well as the evaluation of the new condition number of Q^T. In this case, where a control input exists, the same procedure is followed and, furthermore, the results obtained will be equal to the ones that would be obtained if $u(t)=0$ because the observability matrix Q is independent of $u(t)$.

The case of multi-output with control input is, again, very similar to the multi-output case with $u(t)=0$. Consider equation 4.69 where the matrix Q^T is no longer a nxn matrix. In order to compute the initial states the direct inversion of Q^T cannot be made because Q^T is not a square matrix . Premultiplying equation 4.69 by Q equation 4.74 is obtained

$$Qy_d(t_0)=QQ^Tx(t_0)+Q\Omega u_d(t_0) \tag{4.74}$$

and solving it for $x(t_0)$

$$x(t_0)=(QQ^T)^{-1}y_d(t_0)-(QQ^T)^{-1}\Omega u_d(t_0) \tag{4.75}$$

Observe that in this case the degree of ill-conditioning of QQ^T is the one that has to be studied instead of the degree of ill-conditioning of Q^T. For the evaluation of the relative degree of observability the generalized graphical method is the most suitable one because the matrix involved in the derivation of the algorithm is the QQ^T matrix. For the case of improvement of the relative observability of a particular state variable the problems that were presented in the Section D of this Chapter are also valid here because the square matrix QQ^T is independent of the control input.

F. SUMMARY

In this Chapter a method to improve the relative observability of a weakly observable state was developed for the single output regulator $(u(t)=0)$ case. This method was based on the concepts of spectral factorization of a matrix with distinct eigenvalues discussed in [Ref. 10: p. 609]. The method has the goal of minimizing the condition number of the Q^T matrix. An algorithm was derived and a Fortran program (COBS) presented in Appendix B was written. This program improves the relative degree of observability of the least observable state. Documentation is provided to enable the user to improve the relative degree of observability of a particular state-variable other than the least observable one. This method was then extended for the multi-output case where the concept of the pseudo-inverse of a matrix was used. The factorization of Q^T into the product of three matrices where only one was not entirely dependent on the plant was performed and examples were presented to demonstrate that this general case factorization contains the single output case. The method to improve the relative degree of observability of a particular state-variable was then presented and shown to be different from the single output case as well as much more elaborate. Finally, the case where a control input is present was studied both for the single and multi-output cases, leading to the conclusion that the fact that $u(t) \neq 0$ doesn't change in any way what has been said for the cases where $u(t) = 0$.

V. ANALYSIS OF THE RESULTS

A. INTRODUCTION

In the last two Chapters algorithms have been developed to test the relative degree of observability of a LTI system and to improve the degree of observability of a weakly observable state. As previously mentioned, these algorithms are implemented in two programs (TOBS and COBS) whose performance will be discussed in this Chapter. The results obtained while running the programs will be evaluated and some observations will be presented.

B. PERFORMANCE OF TOBS

This program tests and ranks the relative degree of observability for each state variable using three out the four methods presented in Chapter III, namely the upper-bound error method, the graphical method and the generalized graphical method. In the first phase of testing the program, the performance of the methods was investigated using examples already tested and used in [Ref. 7]. The Table below shows the values for the degrees of observability obtained. As can be seen all the three methods agree on what are the least and most observable states. The matrices A and C of this system are:

$$A = \begin{bmatrix} -12 & 3 \\ 3 & -4 \end{bmatrix} \qquad C = [0 \quad 1]$$

TABLE 1

TOBS OUTPUT FOR ABLIN'S SECOND ORDER EXAMPLE

Upper-Bound Error		Graphical Method		Generalized Graphical	
x_2	1.0	x_2	1.0	x_2	0.68
x_1	0.6	x_1	0.3	x_1	0.05

Table 1 shows that in all the cases x_2 is the most observable state. This makes sense because the matrix C shows that x_2 is the state variable that is being measured. Once TOBS was tuned and running as desired, several state-space representations of transfer functions with different C matrices were studied. The maximum order tested was order ten. As the order of the plant starts increasing the graphical method begins disagreeing with the other two methods in the ranking of the least observable state. However, the results of the upper-bound and generalized graphical methods agree with one another in almost all cases. This behavior can be understood if one recalls the discussion of the graphical method presented in Chapter III. This method relies on finding the largest component of the vector relating the states with the measurements. As the order of the system increases the surfaces representing the relations between the states and the measurements start being planes and hyperplanes where the graphical reasoning behind this simple procedure is no longer easy to apply. However, even for high order systems, the most observable state is almost always the same in all three methods. The big advantage of the graphical method is the small amount of computation required, enabling the user to obtain a quick determination of the relative observability of the system in low order or reduced order models. Furthermore, due to its relatively easy graphical interpretation, it can be quite helpful in understanding the ideas that serve as the basis for this type of method. Although agreeing with the generalized graphical method, the upper-bound error method suffers from the fact that it needs a matrix inversion to obtain the required results. With this method, the evaluation of the relative degree of observability of the state variables cannot be obtained when the system is not completely observable because the inversion of the observability matrix is needed. In any of the other methods this information can still be obtained even if one or more states are not observable. The generalized method seems to be the best one because although it presents a little more difficult graphical interpretation, it doesn't need any matrix inversion and, more importantly, it can be readily used to compute the relative degree of observability for each state variable in the multi-output case while the other two methods, as they are developed, need some changes before they can be used in evaluating the relative degree of observability for each state variable in the multi-output case.

As was previously mentioned, several runs were made using state-space representations of transfer functions from order two to order ten for different combinations of the entries of the C matrix. Since a large amount of computer output would probably not be of much help in understanding the relative performance of the methods described in Chapter III, a summary of the results is given in Table 2.

TABLE 2

SUMMARY OF THE RESULTS OBTAINED USING THREE DIFFERENT METHODS

Order of the Plant	R_{12}	R_{13}	R_{23}
2	100%	100%	100%
3	95%	95%	100%
4	93%	93%	99%
5	88%	87%	97%
6	82%	82%	97%
7	75%	73%	98%
8	71%	71%	96%
9	69%	68%	94%
10	65%	65%	94%

In Table 2 column one contains the order of the plant. Columns two, three and four contain values that are a measure of the correlation between the graphical method (number one) and upper-bound error (number two) -R_{12}-, the graphical and the generalized graphical method (number three) -R_{13}- and finally, between the upper-bound error method and the generalized graphical method -R_{23}-. The values in Table 2 were computed using equation 5.1

$$R_{ij} = \frac{n_m}{n\, n_r} \qquad i, j = 1,2,3 \tag{5.1}$$

where n is the order of the system (number of states), n_r is the number of runs performed for the same value of n but for different entries of the C matrix and n_m is the number of matches between the two methods insofar as the relative degree of observability is concerned. Notice that the product n n_r represents the maximum possible number of agreements. For instance, suppose that the order of the system is 8 and one performs 10 runs. The maximum possible number of matches is 80. However, if the two methods only agree in 60 ($n_m = 60$) out of the 80 possible, the agreement percentage will be 75% and this is the value that appears in the corresponding columns of Table 2. Looking at this Table one can observe what was mentioned earlier in this Chapter, namely the degradation of the graphical method for the evaluation of the relative degree of observability as the order of the system increases. Actually, the disagreement between the graphical method and the other two is concentrated in the least observable states. In most of the cases the most observable states were the same within one level of ranking in all three methods. As can be seen in the fourth column of Table 2, the agreement percentage is slowly decreasing as the order of the system increases. This occurs due the way the values are computed. In fact, the values for the degrees of observability for the most observable states are, in most cases, of the same order of magnitude for high order systems. A slight change in one digit can produce a different ranking in the state-variables, which is interpreted by equation 5.1 as a mismatch, producing a decrease in the agreement percentage. However, as the order of magnitude of the degree of observability is the same, one can consider that the two methods still agree as far as the relative degree of observability is concerned. Of course, this decision can only be made by direct analysis of the computer output. To illustrate this characteristic behavior the following example is presented. Consider the sixth order system where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -6 & -8 & -10 & -4 & -1 & -3 \end{bmatrix} \qquad C = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

The output of TOBS for this system is given in Table 3.

TABLE 3

TOBS OUTPUT FOR A SIXTH ORDER EXAMPLE

Upper-Bound Error		Graphical Method		Generalized Graphical	
x_6	1.0	x_5	1.0	x_6	0.999
x_5	1.0	x_6	1.0	x_5	0.997
x_2	0.66	x_2	0.52	x_2	0.915
x_4	0.53	x_4	0.08	x_4	0.522
x_1	0.51	x_3	0.06	x_3	0.421
x_3	0.26	x_1	0.05	x_1	0.321

As can be seen in columns four and six, although theoretically the upper-bound error and the generalized graphical methods do not agree in the two most observable states, they are, in fact, equally observable because their degree of observability is of the same order of magnitude within the same method. This just shows that one must be careful in arriving at any kind of conclusion just based on the output of the computer program.

The third and probably more interesting part of this study was the test of models of real plants. The first model studied was the reduced order model of a SL-7 hull with three degrees of freedom. The state space representation was derived in [Ref. 18: pp. 585,592] and is shown below :

$$\begin{bmatrix} v \\ r \\ \Psi \end{bmatrix} = \begin{bmatrix} -2.64e\text{-}2 & -8.027 & 0 \\ -1.35e\text{-}4 & -5.87e\text{-}2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ r \\ \Psi \end{bmatrix}$$

where:

v is the sway rate

r is the yaw rate

Ψ is the yaw

Two runs of this model were performed with two different C matrices. The results presented in Table 4, show the output of TOBS for $C = [\,0 \;\; 1 \;\; 1\,]$

TABLE 4

TOBS OUTPUT FOR THE SL-7 HULL MODEL -CASE I-

Upper-Bound Error		Graphical Method		Generalized Graphical	
x_2	1.0	x_2	0.89	x_2	0.99
x_3	0.71	x_3	0.39	x_3	0.98
x_1	0.02	x_1	0.01	x_1	0.01

and Table 5 shows the output of TOBS for $C = [\,0 \;\; 0 \;\; 1\,]$.

As can be seen in Tables 4 and 5, in both cases the three tests gave v, the sway rate, as the least observable state, which agrees quite strongly with one's intuition about the problem. Furthermore, in [Ref. 18] the author mentions that for some type of studies the sway rate can be neglected in a reduced order model because its influence in the behaviour of the third order model is not very strong.

TABLE 5

TOBS OUTPUT FOR THE SL-7 HULL MODEL -CASE II-

Upper-Bound Error		Graphical Method		Generalized Graphical	
x_3	1.0	x_3	1.0	x_3	1.0
x_2	1.0	x_2	1.0	x_2	0.99
x_1	0.01	x_1	0.01	x_1	0.01

The second model studied in this phase is a seventh order plant that is presented as example two of [Ref. 19], where the author is seeking the identification of the essential states to reduce the order of the model. With his test he finds that states x_4 x_5 x_6 and x_7 are the ones that contain more information on the dominant modes of the plant and are, therefore, the ones to keep in the reduced order model. Performing the three tests described in Chapter III on his model, the results shown in Table 6 are obtained. The relative observability tests performed by TOBS show that the most observable states are x_4, x_5, x_6 and x_7 (with special strength in the three last ones), which agrees with the results obtained by the authors of [Ref. 19]. An interesting point to notice in this example is that there is a difference of one order of magnitude in the degrees of observability of x_4, x_1 and x_2 between the graphical and the generalized graphical method. If only the graphical method was available one could conclude that x_4, x_1, x_2 and x_3 were all weakly observable. However, the upper-bound error method gives magnitudes for the degrees of observability of the state variables that are of the same order of magnitude as the ones given by the generalized graphical method. Therefore, in this example only x_3, and perhaps x_2, would be classified as weakly observable states.

C. PERFORMANCE OF COBS

The algorithm developed in Chapter IV for the single output case lead to COBS, a Fortran program, that improves the relative degree of observability of the least

TABLE 6

TOBS OUTPUT FOR THE MODEL OF EXAMPLE 2 OF [REF. 19]

Upper-Bound Error		Graphical Method		Generalized Graphical	
x_6	1.0	x_6	1.0	x_6	0.99
x_5	1.0	x_5	1.0	x_5	0.99
x_7	0.99	x_7	0.79	x_7	0.98
x_4	0.69	x_4	0.08	x_4	0.92
x_1	0.68	x_1	0.05	x_1	0.81
x_2	0.11	x_2	0.04	x_2	0.80
x_3	0.01	x_3	0.01	x_3	0.02

observable state. Several runs were made of plants of order two up to order seven. For a few cases of the seventh order plants the program took a long time to perform the minimization of the condition number of Q^T. By relaxing the minimization stopping criterion the computing time decreased a bit but not dramatically. However, the results that were obtained were very satisfactory because the least observable state became the most or the second most observable state in all the cases with the corresponding mininization of the condition number of Q^T.

An interesting aspect to observe is what happens before and after the minimization from the graphical point of view. Consider, as an example, the second order system presented in [Ref. 7] where initially

$$A = \begin{bmatrix} -12 & 3 \\ 3 & -4 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

and the least observable state is x_1 as shown in Table 1. Before the minimization the condition number of Q^T is 11.667 where

$$Q^T = \begin{bmatrix} 0 & 1 \\ 3 & -4 \end{bmatrix}$$

After running COBS the new C matrix is $C = [\ 1.17 \quad 1\ ]$. Notice that the entry c_{11} was the one changed during minimization because this entry corresponds to the least observable state whose relative observability is to be improved. With this new C matrix the transpose of the observability matrix is

$$Q^T = \begin{bmatrix} 1.17 & 1.0 \\ 11.00 & -0.5 \end{bmatrix}$$

whose condition number is now equal to 2.89. The new ranking of the state variables based on the relative degree of observability is shown in Table 7. Notice again the difference of one order of magnitude between the graphical and generalized graphical methods. The problem is now the inverse of what it was, i.e., now the graphical method agrees with the upper-bound error method in the order of magnitude of the degrees of observability.

TABLE 7

NEW RANKING OF THE STATE VARIABLES FOR ABLIN'S EXAMPLE

Upper-Bound Error		Graphical Method		Generalized Graphical	
x_1	0.91	x_1	0.85	x_1	0.84
x_2	0.65	x_2	0.53	x_2	0.01

The goal was reached, i.e., the state that once was the least observable one is now the most observable and the condition number of Q^T decreased.

In order to obtain a graphical interpretation of what occurs in the system by changing the C matrix, the procedure used in Chapter III Section E can be used again.

Substituting the Q^T matrix before the minimization in equation 3.19 the equations of two lines will be obtained. These two lines can be plotted in the $x_1(t_0)$, $x_2(t_0)$ plane, assuming that y_1 and y_2 are given measured values. For plotting purposes y_1 and y_2 are considered proportional to the two horizontal lines presented at the top of Figures 5.1 and 5.2. Looking at Figure 5.1 it can be seen that the uncertainty region for $x_2(t_0)$ is equal to the uncertainty region produced by the errors in the measurements. If one relates these results with those of Table 1 one can see that they agree because x_2 is the most observable state. Furthermore, two methods give for this state variable a value of 1 for the degree of observability which means that an error in the measurements is projected with the same order of magnitude when the computation of the initial states takes place. The state-variable x_1 is the least observable because its uncertainty region is larger than the one produced by an error in the measurements. Substituting Q^T after minimization in equation 3.19 another different set of equations relating the states and the measurements of the output are obtained. These two lines are presented in Figure 5.2 in the usual $x_1(t_0)$, $x_2(t_0)$ plane. The distances that correspond to y_1 and y_2 are now the double of what they were in Figure 5.1 but their ratio is the same. This rescaling was done to obtain a cleaner plot of the two lines. Notice that now the uncertainty region for $x_{1a}(t_0)$ is of the same order of magnitude as the one produced by the measurements of the output while the uncertainty region for $x_{2a}(t_0)$ increased. The condition number of Q^T decreased because the smaller angle between the lines in Figure 5.1 is about 33° and in Figure 5.2 is about 42°.

In the runs of COBS with the different plants in about 40% of the cases the initially least observable state became the second most observable, but with a value for the degree of observability of the same order of magnitude as the most observable state. Unfortunately, in the real world, the user may not be able to change the C matrix according to the least observable state. Generally there are constraints both for the entries of the C matrix that one is allowed to change as well as for the values that the entries can have. The program COBS allows a change in the entry of the C matrix chosen by the user although the default is the the change in the entry of C associated with the least observable state. As previously mentioned, documentation in the listing of the program is provided to guide the user to set the program to change any other entry in C. Despite the somewhat unrealistic philosophy upon which this algorithm was based, it provides a way of checking how the system and the states behave with respect to one another, which can be helpful to carry out further investigations.

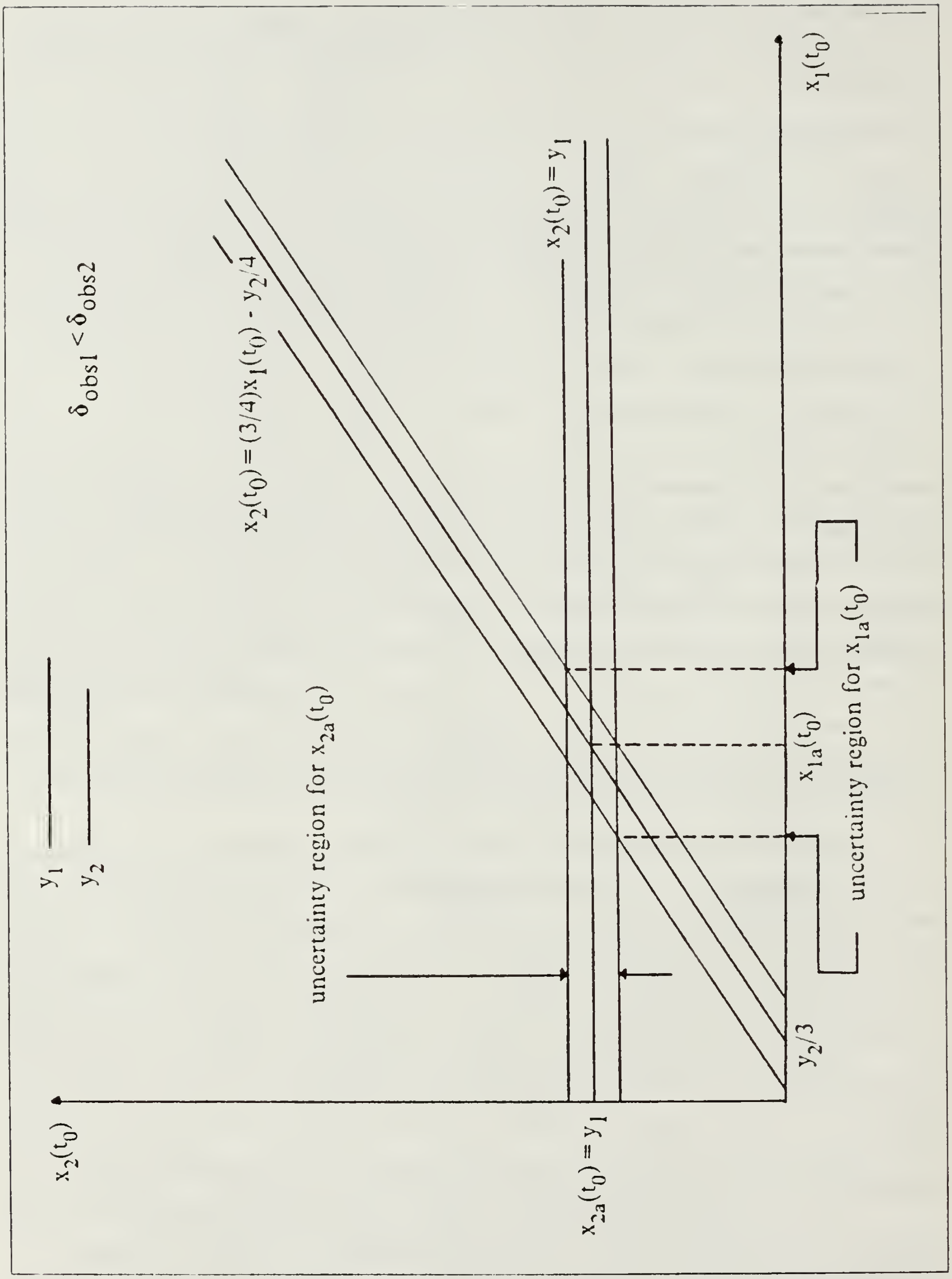
$\delta_{obs1} < \delta_{obs2}$
y_1
y_2
$x_2(t_0) = (3/4)x_1(t_0) - y_2/4$
$x_2(t_0) = y_1$
uncertainty region for $x_{2a}(t_0)$
$x_{1a}(t_0)$
uncertainty region for $x_{1a}(t_0)$
$y_2/3$
$x_2(t_0)$
$x_1(t_0)$
$x_{2a}(t_0) = y_1$

Figure 5.1 Graphical interpretation of Ablin's example before improvement.

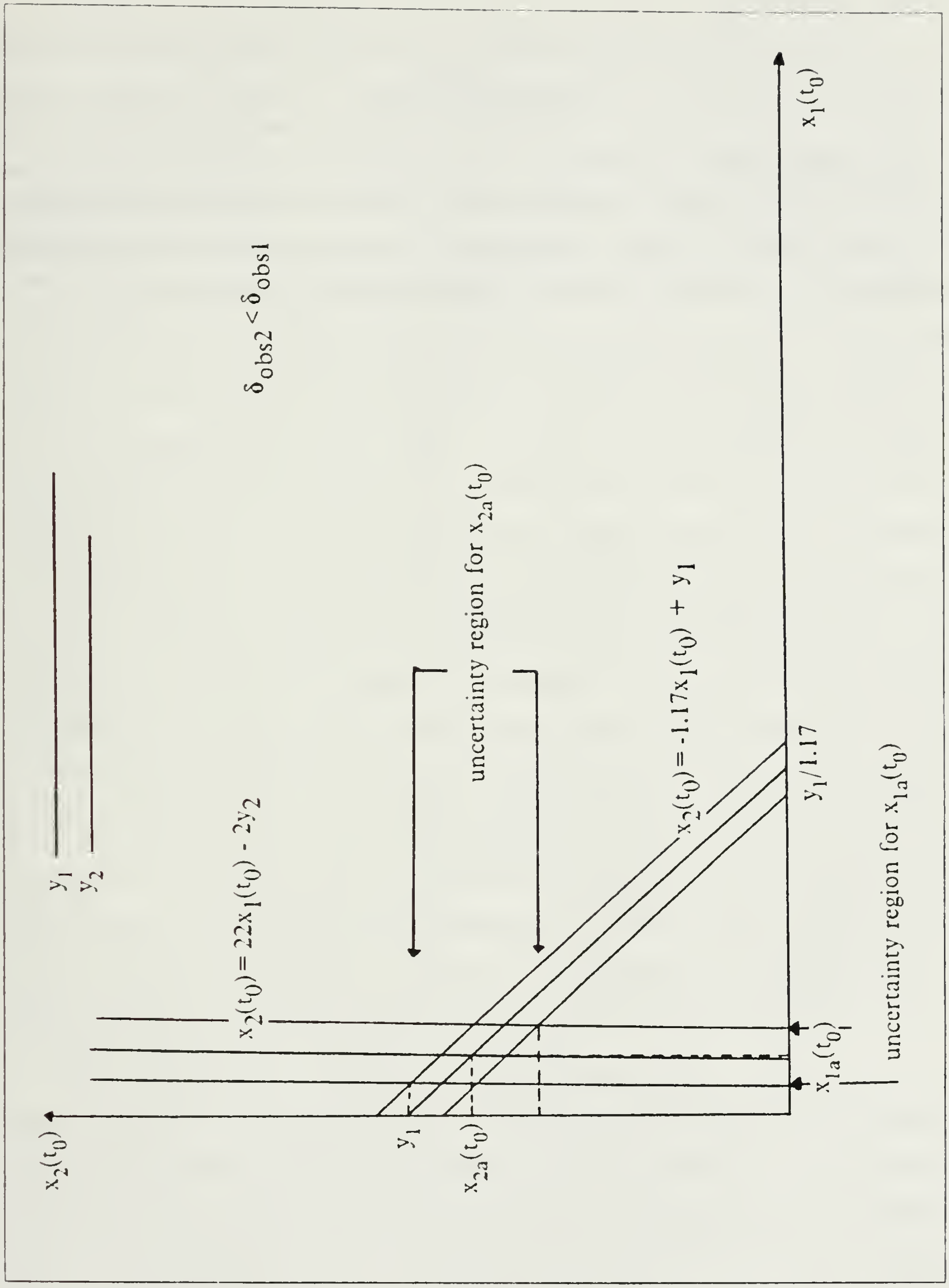
$x_1(t_0)$
$x_2(t_0)$
$\delta_{obs2} < \delta_{obs1}$
y_1
y_2
$x_2(t_0) = 22x_1(t_0) - 2y_2$
uncertainty region for $x_{2a}(t_0)$
$x_2(t_0) = -1.17x_1(t_0) + y_1$
$y_1/1.17$
uncertainty region for $x_{1a}(t_0)$
$x_{1a}(t_0)$
y_1
$x_{2a}(t_0)$

Figure 5.2 Graphical interpretation of Ablin's example after improvement.

D. SUMMARY

In this Chapter the performance of the methods used to evaluate the relative degree of observability of a set of state variables was analyzed and some important results were obtained from the use of the program that performs these tests (TOBS). The analysis of the performance of the algorithm in improving the relative degree of observability of a weakly observable state was done and the graphical interpretation of a second order plant before and after the improvement was studied. Finally the inherent advantages and disadvantages of this algorithm were pointed out.

VI. CONCLUSION

This thesis further develops the issue raised by Brown [Ref. 2] concerning the question of "how observable a linear system is". This issue, which leads to the concept of relative observability, is addressed in four different ways in the thesis. Two of the methods used to test the relative observability of an LTI system are based on graphical interpretations of the relations that exist between the measurements of the output and the states. In the first graphical method the information imbedded in the angles between the vectors whose components are the row entries of the observability matrix and the coordinate axes $(x_1(t_0), x_2(t_0), \ldots, x_n(t_0))$ is used to determine the relative observability of the state variables. The smaller the angle between any vector and an axis of the state-space, the smaller the uncertainty region and the higher the degree of observability for that state-variable. Since the angles between the vectors and the axes are directly related to the magnitudes of their components along these axes, an easy algorithm to check the relative degree of observability of a state-variable was derived, constituting one of the three methods contained in the computer program TOBS. During the study of this method it was observed that a weakly observable state was more sensitive to changes in the magnitudes of the errors in the measurements than a strongly observable state. The second graphical method is called the generalized graphical method and is developed from the idea of searching for the direction of maximum error yielding a minimum for the observability function. This direction turns out to be given by the eigenvector associated with the smallest eigenvalue of the QQ^T matrix. Therefore, the smallest component of the eigenvector associated with the smallest eigenvalue of QQ^T corresponds to the most observable state while the largest component corresponds to the least observable state.

Two other analytical methods were presented and explained: the upper-bound error method and the standard deviation error method. Although these two methods seem to be conceptually different from the two graphical methods, a common thread was the fact that the degree of ill-conditioning of the Q^T matrix plays an important role in the theoretical development of both graphical and analytical methods because a highly ill-conditioned matrix has a twofold meaning. On one hand it means that the rows of that particular matrix have a high degree of dependency (nearly parallel) and on the other hand it means that the determinant of that matrix is a very small number (which causes large entries in the inverse of the matrix). The high degree of dependency among the rows of Q^T is one of the basic ideas behind the development of the graphical method while the small number for the determinant is one of the key ideas behind the understanding of both the upper-bound error and the standard deviation error methods. The graphical, upper-bound error and generalized graphical methods were throughly tested in TOBS using different models. All three methods agreed in almost all cases in ranking the relative observability of the states. However, the graphical method degraded as the order of the system increased. Despite this feature, the graphical method has a big advantage: it requires a very small amount of computation, enabling the user to obtain a quick determination of the relative observability in low order and reduced order system models. Furthermore, due to its relatively easy graphical interpretation it can be very helpful in understanding both the concept of an ill-conditioned matrix and the role that this characteristic of a specific matrix has in computing the relative observability of a system. The upper bound error method is reliable but has two disadvantages: the requirement of matrix inversion for every evaluation of the degree of observability and the impossibility of using the method when the system is not completly observable. The generalized graphical method is the approach whose performance is more satisfactory because it can be used even if the system is not completly observable, doesn't degrade when the order of the system increases and doesn't need any matrix inversion for the required results to be obtained. Furthermore, it can be readily used to compute the relative degree of observability of a multi-output system while the graphical and upper-bound error methods require some changes before being ready to compute the relative degree of observability for each state-variable.

In the second part of this thesis an algorithm to improve the relative degree of observability for the single output regulator case was developed and tested in the computer program COBS. This algorithm was based on the concepts of spectral factorization of a matrix, allowing the decomposition of the Q^T matrix as the product of three matrices, two of them being fully dependent on the the plant and the other depending on both the plant and the C matrix. By altering the entry in the C matrix associated with the least observable state, the minimization of the condition number of Q^T was achieved and the least observable state became the most or the second most observable state in all the cases that were tested. The single-output regulator case was then generalized, first to a multi-output regulator problem and finally to a multi-output servo problem. The use of the concept of the matrix pseudo-inverse was needed for the theoretical development of the multi-output problem and the spectral factorization of Q^T was found to be possible although hard to accomplish. Two illustrative examples were presented. An algorithm to improve the relative degree of observability of a specific state was also proposed, assuming the previous knowledge of what entries in the C matrix could be changed and how much could they be altered. Finally it was shown that the servo problem added nothing new to what had been accomplished for both the single output and the multi-output case because the key matrices (Q^T for the single output and QQ^T for the multi-output) are always independent of the control input.

In conclusion, amongst the methods developed to test the relative degree of observability of the system, it seems that the generalized graphical method is the most reliable one although the graphical method can give quick and easily computed information about the most observable states. Since the upper-bound error method needs a matrix inversion, it may become computationally involved for high order systems and therefore its use should be considered depending on the order of the system to be studied.

By using the algorithms developed in this thesis, the designer can obtain much insight into a complex system with a relatively small amount of computations. Exactly how these ideas may be used will depend on the designer's knowledge of the capabilities of the system as well as the specifications of the problem under study.

There are two recommendations for further study. The first one is the implementation and test of the algorithm proposed for the improvement of the degree of observability of the system for the multi-output case. The second one is the study of the issues addressed in this thesis for linear time-varying systems.

APPENDIX A

TESTING THE DEGREE OF OBSERVABILITY OF AN LTI SYSTEM

```
      PROGRAM TOBS
C
C
C******************* PURPOSE ****************************
C THIS PROGRAM EVALUATES THE DEGREE OF OBSERVABILITY OF *
C A SYSTEM, WHOSE MATRIX A AND C ARE IN A DATA FILE, BY *
C THREE DIFFERENT METHODS: A  GRAPHICAL METHOD,  UPPER  *
C BOUND ERROR METHOD C AND THE GENERALIZED GRAPHICAL    *
C METHOD. THE MAX C ORDER FOR THE PLANT IS ORDER TEN.   *
C THE OUTPUT IS GIVEN C IN A DATA FILE WHERE THE STATES *
C ARE PRESENTED IN DECREASING ORDER OF RELATIVE OBSERV. *
C INPUT DATA IS FILE FT04F001 - LOGICAL RECORD IS 80    *
C OUTPUT DATA IS FILE FT08F001 - LOGICAL RECORD IS 133  *
C********************************************************
C
C
C***** DECLARATION OF VARIABLES *****
C
      IMPLICIT REAL*4 (A-H,O-X)
      DIMENSION A(50,50),QT(50,50),STATE(50),TEST(50),DGOBS(50)
      DIMENSION DGTEST(50),C(1,50),QTN(50,50),QTNI(50,50),FLAG(50)
      DIMENSION WKAREA(130),QN(50,50),QQN(50,50)
      COMPLEX W(50),Z(50,50)
C
C
C***** READ IN ORDER OF SYSTEM N, AND MATRICES A AND C ******
C
      READ(4,*)N
C         WRITE(6,600)I
          READ(4,*)(C(1,I),I=1,N)
      DO 20 I=1,N
C              WRITE(6,700)I,J
               READ(4,*)(A(I,J),J=1,N)
 20   CONTINUE
C
C
C***** OUTPUT MATRICES A AND C *****
C
      WRITE(8,450)
      WRITE(8,350)(C(1,I),I=1,N)
      WRITE(8,550)
      DO 25 I=1,N
          WRITE(8,350)(A(I,J),J=1,N)
 25   CONTINUE
C
C
C***** COMPUTE Q TRANSPOSE ******
C
      DO 30 J=1,N
          QT(1,J)=C(1,J)
 30   CONTINUE
C
      DO 60 I=2,N
          DO 50 JL=1,N
               QT(I,JL)=0.
               DO 40 JC=1,N
                    QT(I,JL)=QT(I,JL)+QT(I-1,JC)*A(JC,JL)
 40            CONTINUE
 50       CONTINUE
 60   CONTINUE
C
```

```
C
C***** NORMALIZE Q TRANSPOSE BY ROW *****
C
      DO 80 I=1,N
          SUM=0.
          DO 70 J=1,N
               SUM=SUM+QT(I,J)*QT(I,J)
 70       CONTINUE
          IF(SUM.EQ.0.)THEN
               QTN(I,JC)=QT(I,JC)
               QN(JC,I)=QTN(I,JC)
          ELSE
               DO 75 JC=1,N
                    QTN(I,JC)=QT(I,JC)/SQRT(SUM)
                    QN(JC,I)=QTN(I,JC)
 75            CONTINUE
          END IF
 80   CONTINUE
C
C
C***** FIND BIGGEST ENTRY IN EACH COLUMN *****
C
      DO 100 J=1,N
          BIG=0.
          DO 90 I=1,N
               IF(ABS(QTN(I,J)).GT.BIG)THEN
                    BIG=ABS(QTN(I,J))
               END IF
 90       CONTINUE
          FLAG(J)=1.
          IF(BIG.EQ.0.)THEN
               WRITE(6,800)J
               WRITE(8,800)J
               FLAG(J)=0.
          END IF
          STATE(J)=BIG
          TEST(J)=STATE(J)
 100  CONTINUE
C
C
C***** COMPUTE THE RELATIVE DEGREE OF OBSERVABILITY ******
C*****     USING THE UPPER-BOUND ERROR METHOD       ******
C
      NSTOP=1.
      DO 105 I=1,N
          NSTOP=NSTOP*FLAG(I)
 105  CONTINUE
      IF(NSTOP.EQ.0.)THEN
          WRITE(6,850)
          WRITE(8,850)
      END IF
      IF(NSTOP.EQ.1.)THEN
          CALL LINV2F(QTN,N,50,QTNI,0,WKAREA,IER)
          DO 120 I=1,N
               DEN=0.
               DO 110 J=1,N
                    DEN=DEN+ABS(QTNI(I,J))
 110           CONTINUE
               DGOBS(I)=1./DEN
               DGTEST(I)=DGOBS(I)
 120      CONTINUE
      END IF
C
C
C***** MULTIPLYING QN BY QTN *****
C
      DO 126 I=1,N
          DO124 JL=1,N
               QQN(I,JL)=0.
               DO 122 JC=1,N
                    QQN(I,JL)=QQN(I,JL)+QN(I,JC)*QTN(JC,JL)
 122           CONTINUE
 124       CONTINUE
 126  CONTINUE
C
C
C***** FINDING THE SMALLEST EIGENVALUE AND CORRESPONDING *****
C***********               EIGENVECTOR              ***********
C
      CALL EIGRF(QQN,N,50,1,W,Z,50,WKAREA,IER)
      SMALL=1000.
      DO 200 I=1,N
          IF(REAL(W(I)).LT.SMALL)THEN
               SMALL=REAL(W(I))
               ISMALL=I
          END IF
 200  CONTINUE
C
C
C***** PRINT IN DECREASING ORDER THE DEGREE OF OBSERVABILITY *****
C*****                FOR EACH STATE VARIABLE                *****
C
      WRITE(8,880)
      DO 140 K=1,N
          BPRIN1=0.
          BPRIN2=0.
          BPRIN3=0.
          DO 130 I=1,N
               IF (TEST(I).GE.BPRIN1)THEN
                    BPRIN1=TEST(I)
                    IBIG1=I
               END IF
               IF (DGTEST(I).GT.BPRIN2.AND.NSTOP.EQ.1.)THEN
                    BPRIN2=DGTEST(I)
                    IBIG2=I
               END IF
               QTEST=(1.-ABS(REAL(Z(I,ISMALL))))
               IF(QTEST.GE.BPRIN3)THEN
                    BPRIN3=QTEST
                    IBIG3=I
               END IF
 130      CONTINUE
C
C
C**************** OUTPUT ******************
C
          IF(NSTOP.EQ.0.)THEN
               WRITE(8,900)IBIG1,BPRIN1,IBIG3,BPRIN3
          END IF
          IF(NSTOP.EQ.1.)THEN
               WRITE(8,910)IBIG1,BPRIN1,IBIG2,BPRIN2,IBIG3,BPRIN3
          END IF
          TEST(IBIG1)=-1.
          DGTEST(IBIG2)=-1.
          Z(IBIG3,ISMALL)=CMPLX(-1.,0.)
 140  CONTINUE
C
C
C ***** FORMATS *****
C
 350  FORMAT(' ',F13.7,7(2X,F13.7))
 370  FORMAT(' ',F13.7,7(4X,F13.7))
 450  FORMAT(' ',//,' THE ENTRIES OF MATRIX C ARE :')
 550  FORMAT(' ',//,' THE ENTRIES OF MATRIX A ARE :')
 560  FORMAT(' ',//,' Q TRANSPOSE IS :')
 570  FORMAT(' ',//,' Q TRANSPOSE NORMALIZED IS  :')
 800  FORMAT(' ',//,'STATE X',I2,1X,'IS NOT OBSERVABLE',//)
 850  FORMAT(' ABLIN','S CRITERION CAN NOT BE EVALUATED, ',/,' BECAUSE
    &Q TRANSPOSE IS SINGULAR',/)
880  FORMAT(' ',/,' RELATIVE DEGREE OF OBS. FOR GRAPHICAL ABLIN',' AND
    & GENERALIZED METHODS')
900  FORMAT(' ',2X,'X',I2,2X,F9.6,2X,12('.'),2X,'X',I2,F9.6)
910  FORMAT(' ',2X,'X',I2,2X,F9.6,2X,'X',I2,2X,F9.6,2X,'X',I2,2X,F9.6)
     END
```

APPENDIX B

IMPROVING THE DEG. OF OBS. OF A WEAKLY OBSERVABLE SYSTEM

```
      PROGRAM COBS
C
C
C********************** PURPOSE **********************************
C THIS PROGRAM COMPUTES THE RELATIVE DEGREE OF OBSERVABILITY     *
C OF THE STATE VARIABLES OF A GIVEN SYSTEM BY THREE DIFFERENT    *
C METHODS AND THEN COMPUTES THE NEW ENTRY IN THE C MATRIX IN     *
C ORDER FOR THE DEGREE OF ILL-CONDITIONING OF THE OBSERVABILITY  *
C MATRIX TO BE MINIMUM. IN THIS PROGRAM THE ENTRY OF C TO BE     *
C CHANGED IS THE ONE CORRESPONDING TO THE LEAST OBSERVABLE STATE.*
C IF THE USER WANTS OR NEEDS TO CHANGE OTHER ENTRY OF C RATHER   *
C THAN THE ONE ASSOCIATED WITH THE LEAST OBSERVABLE STATE, THE   *
C INTEGER VARIABLE LR NEEDS TO BE SET FOR THE ORDER OF THE STATE *
C TO BE CHANGED. FOR INSTANCE, IF X4 NEEDS TO BE CHANGED INSTEAD *
C OF THE ENTRY ASSOCIATED WITH THE LEAST OBSERVABLE STATE, THE   *
C USER SHOULD PUT LR=4                                           *
C ****************************************************************
C
C
C***** DECLARATION OF VARIABLES *****
C
      IMPLICIT REAL*4 (A-H,O-X)
      DIMENSION A(10,10),QT(10,10),STATE(10),TEST(10),DGOBS(10)
      DIMENSION DGTEST(10),C(1,10),QTN(10,10),QTNI(10,10),FLAG(10)
      DIMENSION WK0(130),WK1(130),WK2(130)
      DIMENSION QN(10,10),QQN(10,10),P(10,10),M(10,10)
      DIMENSION V(10,10),G(10,10),DA(10,10),DAR(10,10),QTI(10,10)
      DIMENSION AL(1,100),DARI(10,10)
      COMPLEX W(10),Z(10,10)
C
C
C***** READ IN ORDER OF SYSTEM N, AND MATRICES A AND C ******
C
      NT=0
      DEL=0.001
      READ(4,*)N
          READ(4,*)(C(1,I),I=1,N)
      DO 20 I=1,N
               READ(4,*)(A(I,J),J=1,N)
 20   CONTINUE
C
C
C***** OUTPUT MATRICES A AND C *****
C
      WRITE(8,450)
      WRITE(8,350)(C(1,I),I=1,N)
      WRITE(8,550)
      DO 25 I=1,N
          WRITE(8,350)(A(I,J),J=1,N)
 25   CONTINUE
C
C
C***** COMPUTE Q TRANSPOSE ******
C
      CALL MULT1(C,A,N,QT)
C
C
C***** NORMALIZE Q TRANSPOSE BY ROW *****
C
      WRITE(6,700)
 700  FORMAT(' ',' BEGINING COMPUTING THE DEGREES OF OBSERVABILITY')
 26    CONTINUE
       DO 80 I=1,N
           SUM=0.
           DO 70 J=1,N
                SUM=SUM+QT(I,J)*QT(I,J)
 70        CONTINUE
           IF(SUM.EQ.0.)THEN
                QTN(I,JC)=QT(I,JC)
                QN(JC,I)=QTN(I,JC)
           ELSE
                DO 75 JC=1,N
                     QTN(I,JC)=QT(I,JC)/SQRT(SUM)
                     QN(JC,I)=QTN(I,JC)
 75             CONTINUE
           END IF
 80    CONTINUE
C
C
C***** FIND LARGEST ENTRY IN EACH COLUMN *****
C
       DO 100 J=1,N
           BIG=0.
           DO 90 I=1,N
                IF(ABS(QTN(I,J)).GT.BIG)THEN
                     BIG=ABS(QTN(I,J))
                END IF
 90        CONTINUE
           FLAG(J)=1.
           IF(BIG.EQ.0.)THEN
                WRITE(6,800)J
                WRITE(8,800)J
                FLAG(J)=0.
           END IF
           STATE(J)=BIG
           TEST(J)=STATE(J)
 100   CONTINUE
C
C
C***** COMPUTE THE RELATIVE DEGREE OF OBSERVABILITY ******
C*****     USING THE UPPER-BOUND ERROR METHOD       ******
C
       NSTOP=1.
       DO 105 I=1,N
           NSTOP=NSTOP*FLAG(I)
 105   CONTINUE
       IF(NSTOP.EQ.0.)THEN
           WRITE(6,850)
           WRITE(8,850)
       END IF
       IF(NSTOP.EQ.1.)THEN
           CALL LINV2F(QTN,N,10,QTNI,2,WKO,IER)
           DO 120 I=1,N
                DEN=0.
                DO 110 J=1,N
                     DEN=DEN+ABS(QTNI(I,J))
 110            CONTINUE
                DGOBS(I)=1./DEN
                DGTEST(I)=DGOBS(I)
 120       CONTINUE
       END IF
C
C
C***** MULTIPLYING QN BY QTN *****
C
       DO 126 I=1,N
           DO 124 JL=1,N
                QQN(I,JL)=0.
                DO 122 JC=1,N
                     QQN(I,JL)=QQN(I,JL)+QN(I,JC)*QTN(JC,JL)
 122            CONTINUE
 124        CONTINUE
 126    CONTINUE
C
C
C***** FINDING THE SMALLEST EIGENVALUE AND CORRESPONDING *****
C***********               EIGENVECTOR               ***********
C
        CALL EIGRF(QQN,N,10,1,W,Z,10,WK0,IER)
        SMALL=1000.
        DO 200 I=1,N
            IF(REAL(W(I)).LT.SMALL)THEN
                 SMALL=REAL(W(I))
                 ISMALL=I
            END IF
 200    CONTINUE
C
C
C***** PRINT IN DECREASING ORDER THE DEGREE OF OBSERVABILITY *****
C*****                FOR EACH STATE VARIABLE                *****
C
        WRITE(8,880)
        DO 140 K=1,N
            BPRIN1=0.
            BPRIN2=0.
            BPRIN3=0.
            DO 130 I=1,N
                 IF (TEST(I).GE.BPRIN1)THEN
                      BPRIN1=TEST(I)
                      IBIG1=I
                 END IF
                 IF (DGTEST(I).GT.BPRIN2.AND.NSTOP.EQ.1.)THEN
                      BPRIN2=DGTEST(I)
                      IBIG2=I
                 END IF
                 QTEST=(1.-ABS(REAL(Z(I,ISMALL))))
                 IF(QTEST.GE.BPRIN3)THEN
                      BPRIN3=QTEST
                      IBIG3=I
                 END IF
 130        CONTINUE
C
C
C**************** OUTPUT *****************
C
            IF(NSTOP.EQ.0.)THEN
                 WRITE(8,900)IBIG1,BPRIN1,IBIG3,BPRIN3
            END IF
            IF(NSTOP.EQ.1.)THEN
                 WRITE(8,910)IBIG1,BPRIN1,IBIG2,BPRIN2,IBIG3,BPRIN3
            END IF
C
C*************************************************************
C***** THE FOLLOWING THREE LINES OF CODE ARE THE ONES WHERE *****
C***** THE INTEGER VARIABLE LR IS ASSIGNED THE ORDER OF THE *****
C***** LEAST OBSERVABLE STATE. IF THE USER WANTS TO CHANGE  *****
C***** OTHER ENTRY OF C RATHER THAN THE DEFAULT ONE THESE 3 *****
C***** LINES SHOULD BE COMMENTED AND THE 4TH ONE ACTIVATED  *****
C*************************************************************
C
            IF(K.EQ.N)THEN
                 LR=IBIG3
            END IF
C           LR=ORDER OF THE STATE TO BE CHANGED
C
            TEST(IBIG1)=-1.
            DGTEST(IBIG2)=-1.
            Z(IBIG3,ISMALL)=CMPLX(-1.,0.)
 140    CONTINUE
```

```
      CALL CONDN(QT,N,N,CNQ)
      WRITE(8,610)CNQ
      WRITE(6,705)
 705  FORMAT(' ',' FINISHING THE DEGREES OF OBSERVABILITY')
      IF(NT.EQ.1)GO TO 999
C
C
C***** COMPUTE THE LEFT EIGENVECTORS OF A ******
C
      CALL EIGRF(A,N,10,1,W,Z,10,WK1,IER)
C
C
C***** TRANSFORM W IN P WHERE P ARE THE REAL PARTS OF W *****
C
      DO 1000 I=1,N
          DO 990 J=1,N
               P(I,J)=REAL(Z(I,J))
 990      CONTINUE
 1000 CONTINUE
C
C
C***** FIND THE MINIMUM CONDITION NUMBER FOR Q TRANSPOSED *****
C******** AND CORRESPONDING NEW ENTRY IN THE C MATRIX *********
C
      WRITE(6,710)
 710  FORMAT(' ',' MINIMIZING THE CONDITION NUMBER OF QT')
      Y1=CNQ
      CALL MULT2(C,P,N,G,AL,LS,LG)
      LSOLD=LS
      LGOLD=LG
      CALL CONDN(G,N,N,CNG)
      X1=CNG
C
C
C****** COMPUTE THE CNQ FOR A CNG THAT EQUALS ONE *****
C
      PK=1.
      CALL CNEW(C,P,LR,LS,LG,PK,N,CN)
      C(1,LR)=CN
      CALL MULT1(C,A,N,QT)
      CALL CONDN(QT,N,N,CNQ)
      Y2=CNQ
      X2=PK
C
 280  CONTINUE
C
C
C***** FIND THE ORDER OF THE LARGEST AND SMALLEST ENTRY IN G *****
C
      CALL MULT2(C,P,N,G,AL,LS,LG)
C
C
C********** BISSECT THE INTERVAL  X1,X2  *********
C
      X3=(X1+X2)/2.
      CNG=X3
C
C
C***** FIND NEW ORDINATE (CNQ) FOR THE PREVIOUS ABCISSA (CNG) *****
C
      PK=1./CNG
      CALL CNEW(C,P,LR,LSOLD,LGOLD,PK,N,CN)
      C(1,LR)=CN
      CALL MULT1(C,A,N,QT)
      CALL CONDN(QT,N,N,CNQ)
      Y3=CNQ
C
C
C***** COMPUTE THE SLOPE NEAR X3 *****
```

```
C
      XDEL=X3+DEL
      CNG=XDEL
      PK=1./CNG
      CALL CNEW(C,P,LR,LSOLD,LGOLD,PK,N,CN)
      C(1,LR)=CN
      CALL MULT1(C,A,N,QT)
      CALL CONDN(QT,N,N,CNQ)
      YDEL=CNQ
      SLOPE=(Y3-YDEL)/(X3-XDEL)
C
C
C***** RENAME THE NEW POINTS SUCH THAT IN THE INTERVAL *****
C*****    TOWARDS THE DIRECTION OF DESCENDING SLOPE    *****
C*****  THE POINT WITH THE LARGEST ORDINATE IS X1,Y1   *****
C***** AND THE ONE WITH THE SMALLEST ORDINATE IS X2,Y2 *****
C
      IF(SLOPE.GE.0.)THEN
          IF(Y2.LT.Y1)THEN
               IF(Y3.LT.Y2)THEN
                    Y1=Y2
                    X1=X2
                    Y2=Y3
                    X2=X3
               ELSE
                    Y1=Y3
                    X1=X3
               END IF
          ELSE
               IF(Y3.LT.Y1)THEN
                    Y2=Y3
                    X2=X3
               ELSE
                    Y2=Y1
                    X2=X1
                    Y1=Y3
                    X1=X3
               END IF
          END IF
      ELSE
          IF(X2.LT.X1)THEN
               IF(Y3.LT.Y1)THEN
                    Y2=Y3
                    X2=X3
               ELSE
                    Y2=Y1
                    X2=X1
                    Y1=Y3
                    X1=X3
               END IF
          ELSE
               IF(Y3.LT.Y2)THEN
                    Y1=Y2
                    X1=X2
                    Y2=Y3
                    X2=X3
               ELSE
                    Y1=Y3
                    X1=X3
               END IF
          END IF
      END IF
      IF(0.5*ABS(X2-X1).GT.1.E-5)GO TO 280
      WRITE(6,720)
 720  FORMAT(' ',' OPTIMIZATION COMPLETED')
      WRITE(6,725)
 725  FORMAT(' ',' COMPUTING THE NEW DEGREES OF OBSERVABILITY')
C
C
```

```
C*************** NEW OUTUP ******************
C
      WRITE(8,500)
      WRITE(8,350)(C(1,J),J=1,N)
C
      WRITE(8,600)
      L13=13
      NT=NT+1
      IF(NT.EQ.1) GOTO 26
C
C
C ***** FORMATS *****
C
 350  FORMAT(' ',F13.7,7(2X,F13.7))
 370  FORMAT(' ',F13.7,7(4X,F13.7))
 450  FORMAT(' ',//,' THE ENTRIES OF MATRIX C ARE :')
 500  FORMAT(' ',//,' THE NEW ENTRIES OF MATRIX C ARE :')
 550  FORMAT(' ',//,' THE ENTRIES OF MATRIX A ARE :')
 560  FORMAT(' ',//,' Q TRANSPOSE IS :')
 570  FORMAT(' ',//,' Q TRANSPOSE NORMALIZED IS :')
 600  FORMAT(' ',//,' THE NEW DEGREES OF OBSERVABILITY ARE :')
 610  FORMAT(' ',//,' THE CONDITION NUMBER OF Q TRANSPOSED IS EQUAL TO
     &',F13.7)
 800  FORMAT(' ',//,'STATE X',I2,1X,'IS NOT OBSERVABLE',//)
 850  FORMAT(' ABLIN','S CRITERION CAN NOT BE EVALUATED, ',/,' BECAUSE
     &Q TRANSPOSE IS SINGULAR',/)
 880  FORMAT(' ',/,' RELATIVE DEGREE OF OBS. FOR GRAPHICAL ABLIN',' AND
     & GENERALIZED METHODS')
 900  FORMAT(' ',2X,'X',I2,2X,F9.6,2X,12('.'),2X,'X',I2,F9.6)
 910  FORMAT(' ',2X,'X',I2,2X,F9.6,2X,'X',I2,2X,F9.6,2X,'X',I2,2X,F9.6)
 999  WRITE(6,730)
 730  FORMAT(' ',' PROGRAM COMPLETED')
      STOP
      END
c
c
c****************** SUBROUTINES *******************
C
C****** SUBROUTINE TO COMPUTE THE COLUMN NORMS *****
C
C***************** PURPOSE ***********************
C*  THE PURPOSE OF THIS SUBROUTINE IS TO COMPUTE   *
C*THE MAXIMUM SUM OF THE COLUMNS OF A GIVEN MATRIX *
C*************************************************
c
      SUBROUTINE NORM(DA,NL,NC,DBIG)
          REAL*4 DA(10,10)
          DBIG=-1.
          DO 3 J=1,NC
               TOC=0.
               DO 1 I=1,NL
                    TOC=TOC+ABS(DA(I,J))
 1             CONTINUE
               IF(TOC.GT.DBIG)THEN
                    DBIG=TOC
               END IF
 3        CONTINUE
      RETURN
      END
c
c
C***** SUBROUTINE MULT1 (COMPUTES Q TRANSPOSED) ****
C
C***************** PURPOSE ***********************
C*  THE PURPOSE OF THIS SUBROUTINE IS TO COMPUTE   *
C*THE PRODUCT OF THE C AND A MATRICES TO PRODUCE QT*
C*************************************************
c
      SUBROUTINE MULT1(C,A,N,QT)
      REAL*4 C(1,10),A(10,10),QT(10,10)
      DO 30 J=1,N
          QT(1,J)=C(1,J)
 30   CONTINUE
C
      DO 60 I=2,N
          DO 50 JL=1,N
               QT(I,JL)=0.
               DO 40 JC=1,N
                    QT(I,JL)=QT(I,JL)+QT(I-1,JC)*A(JC,JL)
 40            CONTINUE
 50       CONTINUE
 60   CONTINUE
      RETURN
      END
c
c
C**** SUBROUTINE MULT2 (COMPUTES THE MATRIX G) AND GIVES THE ORDER ****
C********* OF THE LARGEST AND SMALLEST ENTRY IN THE G MATRIX **********
C
C****************** PURPOSE ************************
C*  THE PURPOSE OF THIS SUBROUTINE IS TO COMPUTE   *
C*THE PRODUCT OF THE C AND P MATRICES TO PRODUCE G *
C* IT ALSO PROVIDES THE ORDER OF THE SMALLEST AND  *
C*  LARGEST ENTRIES IN THE G MATRIX AS WELL AS AL  *
C*  WHICH IS A ROW VECTOR WHOSE ELEMENTS ARE THE   *
C*                ENTRIES OF G                     *
C**************************************************
c
      SUBROUTINE MULT2(C,P,N,G,AL,LS,LG)
      REAL*4 C(1,10),P(10,10),AL(1,10),G(10,10)
      DO 1090 JL=1,N
          AL(1,JL)=0.
          DO 1080 JC=1,N
               AL(1,JL)=AL(1,JL)+C(1,JC)*P(JC,JL)
 1080     CONTINUE
 1090 CONTINUE
      GRL=-1.E+4
      PQL=1.E+4
      DO 2020 J=1,N
          IF(AL(1,J).GT.GRL)THEN
               GRL=AL(1,J)
               LG=J
          END IF
          IF(AL(1,J).LT.PQL)THEN
               PQL=AL(1,J)
               LS=J
          END IF
 2020 CONTINUE
      DO 1094 I=1,N
          DO 1092 J=1,N
               IF(I.EQ.J)THEN
                    G(I,J)=AL(1,J)
               ELSE
                    G(I,J)=0.
               END IF
 1092     CONTINUE
 1094 CONTINUE
      RETURN
      END
c
c
c**** SUBROUTINE CONDN (COMPUTES THE CONDITION NUMBER OF AN ARRAY) ****
C
      SUBROUTINE CONDN (DAR,NL,NC,DCN)
      REAL*4 DAR(NL,NC),DARI(10,10),WK2(130)
      CALL NORM(DAR,NL,NC,BG)
      CALL LINV2F(DAR,NL,10,DARI,2,WK2,IER)
      CALL NORM(DARI,NL,NC,BGI)
      DCN=BG*BGI
      RETURN
      END
C
C
C***** SUBROUTINE TO COMPUTE THE NEW ENTRY IN THE C MATRIX *****
C
      SUBROUTINE CNEW(C,P,LR,LS,LG,PK,N,CN)
      REAL*4 C(1,10),P(10,10)
      ADD=0.
      DO 2040 I=1,N
          IF(I.NE.LR)THEN
               ADD=ADD+(PK*P(I,LG)-P(I,LS))*C(1,I)
          END IF
 2040 CONTINUE
      CN=(-ADD)/(PK*P(LR,LG)-P(LR,LS))
      RETURN
      END
```

LIST OF REFERENCES

1. R. Kalman, On the General Theory of Control Systems, *Automatic and Remote Control*, pp 481-492, Butterworth and Company, London, (1961).

2. R. Brown, Not Just Observable But How Observable?, *National Electronics Conference Proceedings*, Vol. 22, 1966, pp 709-714.

3. L. Silverman and S. Shokoohi, Linear Time Variable Systems: Balancing and Model Reduction, *IEEE Transactions on Automatic Control*, Vol. AC-28, No. 8, August 1983, pp 810-822.

4. A. Jonckheere and L. Silverman, A New Set of Invariants for Linear Systems-Applications to Reduced Order Compensator Design, *IEEE Transactions on Automatic Control*, Vol. AC-28, No. 10, October 1983, pp 953-964.

5. L. Pernebo and L. Silverman, Model Reduction Via Balanced State Representation, *IEEE Transactions on Automatic Control*, Vol. AC-27, No. 2, April 1982, pp 382-387.

6. H. Elliott and W. Wolovich, A Frequency Domain Model Reduction Procedure, *Automatica*, Vol. 16, 1980, pp 167-177.

7. H. Ablin, *Criterion for Degree of Observability in a Control System*, Chap. 2, Iowa State University of Science and Technology, Ames (1967).

8. K. Ogata, *Modern Control Engineering*, Chap. 16, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1984).

9. K. Astrom and B. Wittenmark, *Computer Controlled Systems-Theory and Design*, Chap. 5, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1984).

10. T. Kailath, *Linear Systems*, Chap. 2, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1980).

11. C. Helstrom, *Probability and Stochastic Processes for Engineers*, Chap. 2, Macmillan Publishing Company, New York (1984).

12. G. Birkhoff and S. Maclane, *A Survey in Modern Algebra*, Eighth Edition, Chap. 7, Macmillan Publishing Company, New York (1953).

13. R. Bellman, *Introduction to Matrix Analysis*, Chap. 7, McGraw-Hill Book Company, Inc., New York (1960).

14. P. DeRusso, J. Roy and M. Close, *State Variables for Engineers*, Chap 4, John Wiley and Sons, Inc., New York (1965).

15. F. Gerald and O. Weatley, *Applied Numerical Analysis*, Third Edition, Chap. 2, Addison-Wesley Publishing Company, Menlo Park California (1984).

16. N. Gavurin, Ill Conditioned Systems of Linear Algebraic Equations, *International Journal of Computer Mathematics*, Vol. 1, No. 1, 1964, pp 36-50.

17. L. Zadeh and C. Desoer, *Linear Systems Theory the State-Space Approach*, Chap. 11, MacGraw-Hill Book Company, Inc., New York (1963).

18. J. Comstock and P. Mandel, *Principles of Naval Architecture*, Chap. 12, SNAME, New York (1967).

19. A. Johnson and W. Daniels, Identification of Essential States for Reduced-Order Models using a Modal Analysis, *IEE Proceedings*, Vol. 132, Pt. D, No. 3, May 1985, pp 111-117.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library,Code 0142 Naval Postgraduate School Monterey, California 93943-5002	2
3.	Professor Donald E. Kirk, Code 62Ki Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, California 93943	2
4.	Professor Roberto Cristi, Code 62Cx Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, California 93943	1
5.	Cpt. Paulo Roberto de Souza Centro Tecnico Aeroespacial - IAE - ESB 12225 S. Jose dos Campos - SP Brazil	1
6.	Direccao do Servico de Instrucao e Treino Edificio da Administracao Central de Marinha 1188 Lisboa-Codex Portugal	1
7.	Lt. Antonio J. Gameiro Marques Direccao do Servico de Instrucao e Treino Edificio da Administracao Central de Marinha 1188 Lisboa-Codex Portugal	2

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CALIFORNIA 93943 5002

221099

Thesis
M351 Marques
c.1 On the relative observability of a linear system.

221099

Thesis
M351 Marques
c.1 On the relative observability of a linear system.

thesM351
On the relative observability of a linea
3 2768 000 75926 0
DUDLEY KNOX LIBRARY